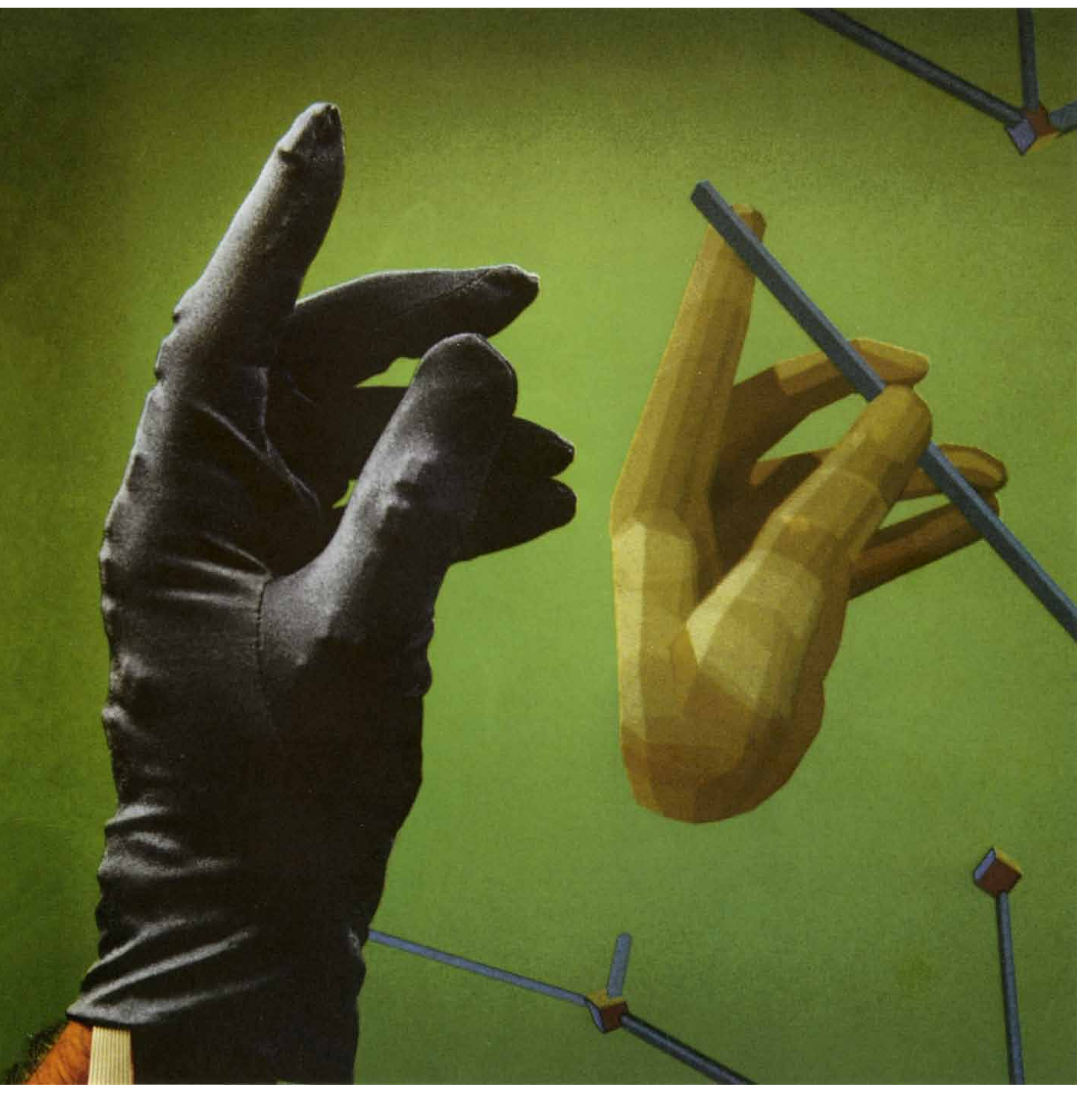
INVESTIGACION Y CIENCIA

Edición en español de **SCIENTIFIC AMERICAN**



AVANCES EN COMPUTACION

Diciembre 1987

575 PTAS.
(IVA INCLUIDO)

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 135
Diciembre 1987

INVESTIGACION Y
CIENCIA *Edición en español de* SCIENTIFIC AMERICAN

12 **LA PROXIMA REVOLUCION INFORMATICA, Abraham Peled**
La revolucion industrial transformó la sociedad. ¿Qué cambiará la revolución informática?

24 **ARQUITECTURAS AVANZADAS DE COMPUTADORES, Geoffrey C. Fox y Paul C. Messina**
Se están proyectando superordenadores que realizarán millones de operaciones por segundo.

34 **MICROCIRCUITOS PARA COMPUTACION AVANZADA, James D. Meindl**
En la carrera para incrementar la velocidad de cálculo, la clave reside en la miniaturización.

52 **PROGRAMAS PARA COMPUTACION AVANZADA, David Gelernter**
La computación en paralelo exige programas que controlen a la vez muchos procesadores.

62 **ALMACENAMIENTO DE DATOS PARA COMPUTACION AVANZADA, Mark H. Kryder**
En cinco años, los dispositivos magnéticos quintuplicarán su capacidad de almacenamiento.

76 **INTERFASES PARA ORDENADORES AVANZADOS, James D. Foley**
Las "entidades artificiales" permiten la comunicación directa entre usuario y ordenador.

86 **AVANCES EN REDES DE ORDENADORES, Robert E. Kahn**
A través de redes comunes, los ordenadores comparten programas, datos y soporte físico.

96 **COMPUTACION AVANZADA EN EL CAMPO DE LA CIENCIA, Piet Hut y Gerald Sussman**
Gracias al ordenador, los astrofísicos conocen las consecuencias de los choques intergalácticos.

108 **COMPUTACION AVANZADA EN EL CAMPO DE LA MEDICINA, Glenn D. Rennels y Edward H. Shortliffe** Los ordenadores intervendrán en los procesos de diagnóstico y tratamiento.

118 **COMPUTACION AVANZADA EN LA INDUSTRIA, Albert M. Erisman y Kenneth W. Neves**
El diseño informático de un avión o un automóvil ahorra dinero y mejora el resultado final.

4 AUTORES
8 HACE...
48 CIENCIA Y SOCIEDAD
128 JUEGOS DE ORDENADOR
136 LIBROS
142 BIBLIOGRAFIA
144 INDICE ANUAL

Colaboradores de este número:

Asesoramiento y traducción:

Luis Bou: *La próxima revolución informática, Microcircuitos para computación avanzada, Programas para computación avanzada* y *Juegos de ordenador*; Gabriel Puig: *Arquitecturas avanzadas de computadores*; Armando Roy Yarza: *Almacenamiento de datos para computación avanzada*; Juan Pedro Adrados: *Interfases para ordenadores avanzados*; Jaime Marcos Escudero: *Avances en redes de ordenadores*; Jordi Ocaña: *Computación avanzada en el campo de la ciencia*; Ana García García: *Computación avanzada en el campo de la medicina*; Josep Enric Llebot: *Computación avanzada en la industria*.

Libros:

Carlos M. Romeo Casabona, David Jou, Juan Gallardo y Rafael Jordana



PORTADA

La ilustración de portada muestra un guante cableado que permite a los usuarios de un ordenador manipular las imágenes en un monitor de gráficos con sus propias manos (véase "Interfases para ordenadores avanzados" por James D. Foley). Por medio de la interfase del guante, el ordenador, de la compañía Symbolics Inc., reconoce los movimientos de los dedos y genera las oportunas sensaciones táctiles, creando la impresión del tacto. Un sensor independiente detecta la posición y orientación de la mano.

Suscripciones:

Prensa Científica, S. A.
Calabria, 235-239
08029 Barcelona (España)
Teléfono 322 05 51 ext. 33-37

Condiciones de suscripción:

España:
Un año (12 números):
4950 pesetas (IVA incluido)

Extranjero:
Un año (12 números): 35 U.S. $

Ejemplar atrasado ordinario:
450 pesetas (IVA incluido)

Ejemplar atrasado extraordinario:
575 pesetas (IVA incluido)

Distribución para España:

Distribuciones de Enlace, S. A.
Bruch, 49 - 08009 Barcelona

Distribución para los restantes países:

Editorial Labor, S. A.
Calabria, 235-239 - 08029 Barcelona

Publicidad:

Madrid:
Gustavo Martínez Ovin
Menorca, 8, bajo, centro, izquierda.
28009 Madrid
Teléfonos 409 70 45 - 409 70 46

Cataluña:
Paulino Muñoz Victoria
Muntaner, 467-469, 3.°, 3.ª
08021 Barcelona
Teléfono 211 92 45

Controlado por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

Fotografía de la portada de Jon Brenneis

Página	Fuente
12	Armando García, IBM Corporation
14	Johnny Johnson
16	IBM Corporation
18-20	James Kilkelly
24	Jon Brenneis
26-32	Jerome Kuhl
35	Jon Brenneis
36	Steve Allen, por cortesía de The Fairchild Semiconductor Corporation
37-44	Hank Iken
52	Animals, Animals/Holt Studios, Ltd.
54-59	Joan Starwood
62	Control Data Corp.
64-65	Ian Worpole
66	W.J. Harley, E.I. du Pont de Nemours & Co., Inc. (*izquierda y centro*); A. Corradi, Magnax, Inc. (*derecha*)
67-70	Ian Worpole
72	IBM Corporation
76	NASA
78	Scott S. Fisher, Douglas E. Kerr, Amy F. Wu y Warren J. Robinett, NASA
79	Steven G. Reznick, Jolen Flores y Jason G. Williams NASA/Ames Research Center
80	Carol Donner
81	James Kilkelly

Página	Fuente
82	Sunil K. Mukerjee y Ali Balali, Universidad George Washington
87	BBN Communication
88-92	George V. Kelvin, Science Graphics
96	James Kilkelly
98-99	Observatorio Nacionál Kitt Peak (*izquierda*) Gabor Kiss (*derecha*)
100	Piet Hut (*arriba derecha*) Gabor Kiss (*arriba izquierda y abajo*)
101	Observatorio Nacional Kitt Peak National
102	Gabor Kiss (*arriba*), Gerald Jay Sussman (*abajo*)
103	Gabor Kiss
109	Larry Fagan y Mark Friss
110	Edward Bell
111	Doris Wells-Papanek y Michael Kahn (*arriba*), Andrew Christie (*abajo*)
112	Cliff Wulfman
113	Randy Miller, Univ. de Pittsburgh
114	Edward Bell
115	Andrew Christie
118 y 123	Boeing Computer Services Co.
120-122	Gabor Kiss
124	Boeing Commercial Airplane Co.
128-130	Andrew Christie

ISSN 0210-136X — Dep. legal: B. 38.999-76

Fotocomposición Tecfa, S.A.
Almogàvers, 189 - 08018 Barcelona

Fotocromos reproducidos por
Reprostudy, S.A.

Gráfica Elzeviriana, S.A.
Nápoles, 249 - Tel. 207 40 11
08013 Barcelona

Printed in Spain - Impreso en España

Los autores

ABRAHAM PELED ("La próxima revolución informática") es vicepresidente de sistemas y director de informática de la división de investigación de IBM. Obtuvo sus primeros títulos universitarios en el Instituto de Tecnología Technion, de Israel, y se doctoró en la Universidad de Princeton. Ha estado en IBM desde 1974, fecha en que ingresó en el centro de investigación Thomas J. Watson. Tras trabajar en otros centros de investigación de IBM, volvió al Watson en 1983, como director de planificación y control de la división de investigación. Peled ocupa el cargo actual desde 1985. Dirige el departamento de ciencias de cómputo del Watson y supervisa las actividades informáticas de los centros y laboratorios de investigación que IBM tiene en todo el mundo.

GEOFFREY C. FOX y PAUL C. MESSINA ("Arquitecturas avanzadas de computadores") están empeñados en un reto: establecer estructuras informáticas fundadas en el procesamiento en paralelo para el trabajo de investigación que se desarrolle en el Instituto de Tecnología de California. Fox, catedrático de física, fue investigador jefe en el programa conjunto del Caltech-Jet Propulsion que habría de conseguir la aplicación de la computación en paralelo al dominio de la ciencia. Messina trabaja en otro proyecto afín: la iniciativa de una supercomputación concurrente, que pretende lograr una computación por procesadores en paralelo cuyo rendimiento no sea inferior al obtenido por los superordenadores al uso. Fox se doctoró en física en la Universidad de Cambridge en 1967; Messina obtuvo ese grado académico por la de Cincinnati en 1972.

JAMES D. MEINDL ("Microcircuitos para computación avanzada") es vicepresidente de asuntos académicos y superintendente del Instituto Politécnico de Rensselaer. Doctor en ingeniería eléctrica por la Universidad Carnegie-Mellon, trabajó en la compañía eléctrica Westinghouse y en el Laboratorio de Electrónica del Ejército. En 1967 ingresó en la Universidad de Stanford para enseñar ingeniería eléctrica. Catedrático desde 1970, fue director de los laboratorios electrónicos de Stanford desde 1972 hasta 1986. Se trasladó al Rensselaer el año pasado.

DAVID GELERNTER ("Programas para computación avanzada") es profesor asociado de informática y ciencias de cómputo en la Universidad de Yale. Tras cursar en ella el primer ciclo de estudios superiores, los prosiguió en la Universidad estatal de Nueva York en Stony Brooks, donde se doctoró en ciencias de cómputo. Gelernter escribe que su interés por los lenguajes de programación y el paralelismo "estuvo en principio motivado por la necesidad de crear instrumentos que permitieran soportar estructuras de programación de nuevo cuño, especialmente en inteligencia artificial. Me atraen también los modelos de memoria".

MARK H. KRYDER ("Almacenamiento de datos para computación avanzada") es profesor de ingeniería eléctrica e informática en la Universidad de Stanford. Allí inició su formación, que terminó en el Instituto de Tecnología de California, donde se recibió de doctor en 1970. Después de permanecer dos años como científico visitante en la Universidad alemana de Regensburg, entró a formar parte de la plantilla del Centro de Investigaciones Thomas J. Watson de IBM. En 1978 pasó a la Universidad de Carnegie-Mellon. Desde 1982 el profesor Kryder ha sido director del centro de tecnología magnética de la universidad.

JAMES D. FOLEY ("Interfases para ordenadores avanzados") enseña informática en la Universidad de George Washington. Está investigando la aplicación de las técnicas de sistemas expertos a las interfases usuario-ordenador. Cursó estudios en la Universidad de Lehigh, continuándolos en la de Michigan, por la que se doctoró en 1969. Foley es coautor de *Fundamentals of Interactive Computer Graphics* y miembro del Instituto de Ingenieros Eléctricos y Electrónicos (IEEE).

ROBERT E. KAHN ("Avances en redes de ordenadores") preside la Corporation for National Research Initiatives, entidad no lucrativa fundada por él mismo para alentar la investigación informática. La Universidad de Princeton le concedió el doctorado en ingeniería eléctrica en 1964. Tras dos años como profesor adjunto en el Instituto de Tecnología de Massachusetts, entró en la firma Bolt, Beranek y Newman, donde colaboró en el desarrollo de la red nacional de ordenadores ARPANET. Pasó luego a un instituto adcrito a la secretaría de Defensa.

PIET HUT y GERALD JAY SUSSMAN ("Computación avanzada en el campo de la ciencia") se dedican a la enseñanza; Hut es profesor de astrofísica en el Instituto de estudios Avanzados y Sussman profesa ingeniería eléctrica en el Instituto de Tecnología de Massachussets. E investigan en colaboración en un proyecto para combinar la astrofísica con la ingeniería en un "observatorio informático". Hut inició su formación en la Universidad de Utrecht y la terminó, con el doctorado, en la de Amsterdam. Sussman recibió su doctorado en matemáticas (1973) en el MIT.

GLENN D. RENNELS y EDWARD H. SHORTLIFFE ("Computación avanzada en el campo de la medicina") investigan integrados en el grupo de ciencias informáticas aplicadas a la medicina de la Universidad de Stanford. Ambos son también médicos en ejercicio en Stanford, Rennels como médico residente del servicio de anestesiología y Shortliffe como profesor de la facultad. Rennels se doctoró en medicina por la Universidad de Dartmouth en 1980; Shortliffe lo hizo por la de Stanford en 1976.

ALBERT M. ERISMAN Y KENNETH W. NEVES ("Computación avanzada en la industria") trabajan en la empresa Boeing de servicios informáticos. En ella, Erisman dirige la división de aplicaciones técnicas y Neves coordina los programas de investigación y desarrollo de la división de servicios científicos. Erisman se doctoró en matemática aplicada por la Universidad de Iowa en 1969, año en que fue contratado por la Boeing. Sus investigaciones se han centrado en la computación científica de alto rendimiento y en algoritmos matriciales que se aplican para analizar los sistemas de energía eléctrica y otros circuitos de gran escala. Erisman ha dado clases también en las universidades de Seattle y Washington. Neves, que obtuvo su doctorado en matemáticas por la Estatal de Arizona, comenzó trabajando en IBM; de ésta pasó a la división de generación de energía nuclear de la Babcock & Wilson Company. En 1975 entró en la Boeing. Entre sus proyectos figura el programa de computación a altísima velocidad.

Hace...

José M.ª López Piñero

...ciento cincuenta años

Murió Tomás Villanova Entraigues, principal introductor de la anatomía comparada en España y figura en la que puede simbolizarse el final de la tradición anatómica iniciada en la Universidad de Valencia a mediados del siglo XVI por los discípulos de Vesalio Pedro Jimeno y Luis Collado.

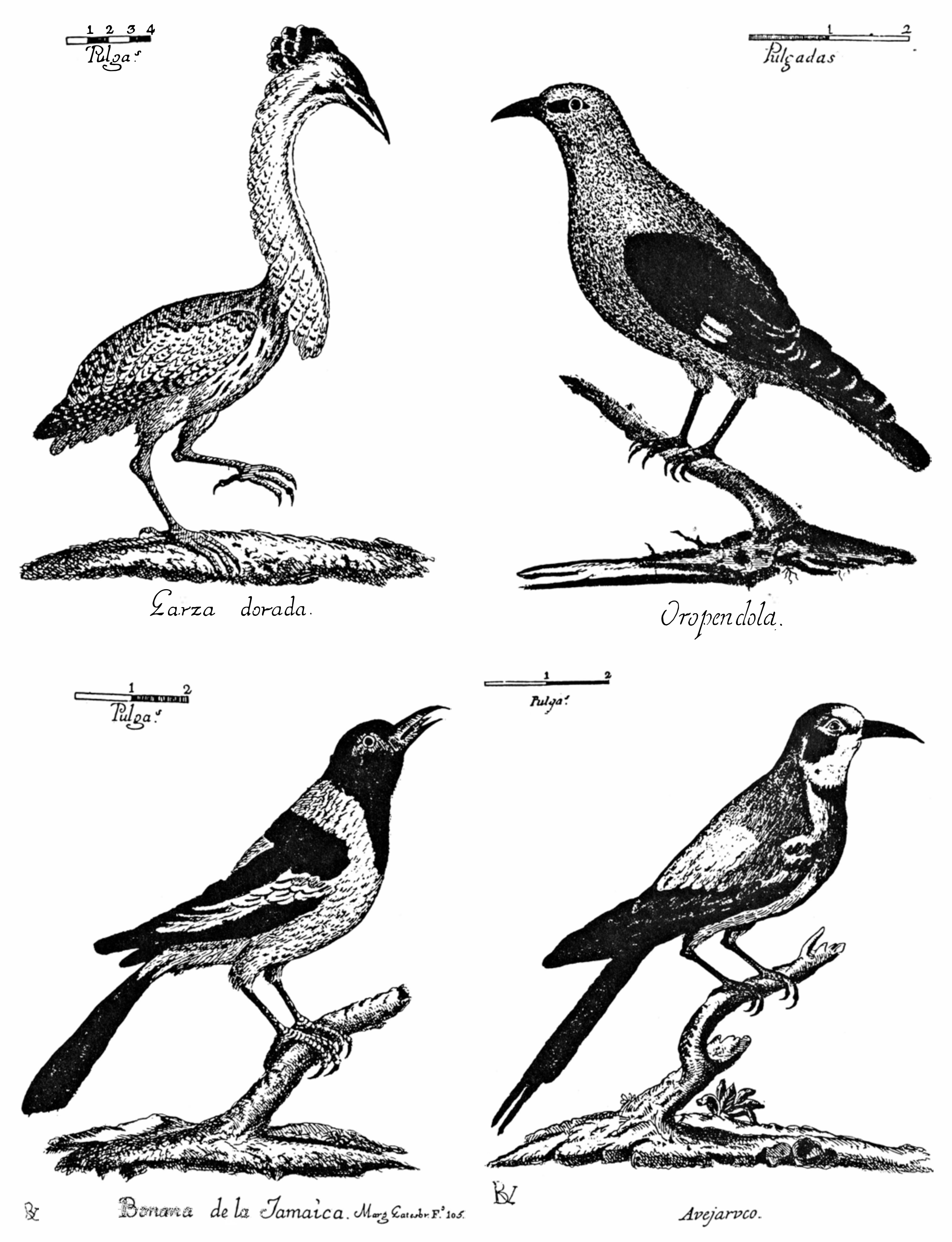
Para encuadrar la actividad de Villanova es necesario tener en cuenta la escasa repercusión que en el terreno morfológico tuvo el decisivo cambio que el célebre plan de estudios del rector Vicente Blasco (1786) significó para el conjunto de la enseñanza médica en la Universidad de Valencia. En contraste con la profunda renovación que dicho plan introdujo en la enseñanza clínica, botánica y química, la reglamentación de la docencia morfológica se limitó a un mero desarrollo de la tradición local de origen renacentista: "Desde Octubre hasta fines de Marzo (el catedrático de anatomía) tendrá la obligación de hacer treinta disecciones en cadáveres, y algunas en animales vivos, para manifestar el sitio y figura de las partes, su composición y enlace, y el uso de cada una para las acciones naturales." En la enseñanza teórica se recomendaba el uso de "las láminas más correctas, y de buenos esqueletos y figuras de cera", consagrando normativamente apoyos didácticos cuya introducción en la Universidad de Valencia se remontaba a Pedro Jimeno. Más importancia tuvo, en cambio, la utilización como libro de texto del *Compendio anatómico*, de Lorenz Heister, típico manual de la Ilustración que alcanzó gran difusión en toda Europa y que había sido traducido al castellano en 1755. El plan de estudios puntualiza, además, que se recomienda "mientras se escribe otro más acomodado que comprenda los nuevos descubrimientos".

Por otra parte, en la implantación efectiva del plan Blasco, la anatomía no contó con un profesor de talla comparable a Félix Miquel, Tomás de Villanova Muñoz y Poyanos, o Vicente Alfonso Lorente, protagonistas de una nueva etapa de la enseñanza clínica, química y botánica. En 1788 falleció Rafael Llombart Choza, infatigable disector, que había sido catedrático de anatomía desde 1752, a continuación de Andrés Piquer. Sus *Theoremata physico-anatomica* demuestran que disponía de una información rigurosa y al día en todos los campos del saber morfológico, desde la anatomía descriptiva y la textural a la embriología, que tenía una mentalidad abiertamente mecanicista y que despreciaba a los "anatomici theoretici... qui sectione cadaverum nunquam occupati fuerunt". Sin embargo, a partir de 1784, la enfermedad que le costaría la vida le impidió ejercer la docencia, por lo cual le sustituyó durante un lustro crucial su hijo Joaquín Llombart Catalá, quien más tarde, a comienzos del siglo XIX, llegaría a ser catedrático de medicina. Llombart Catalá recurrió sin éxito a las oposiciones que en 1790 se convocaron para cubrir la cátedra de anatomía, ya que la plaza fue concedida a Juan Bautista Poeta, que la ocuparía hasta 1801. El resultado de estas oposiciones supera el nivel de lo anecdótico, ya que pesó negativamente en el desarrollo posterior de los saberes morfológicos en Valencia. Llombart Catalá tenía, en efecto, una orientación rigurosa semejante a la de su padre y fue, más tarde, seguidor de las ideas de Xavier Bichat. Por el contrario, Juan Bautista Poeta era un profesor adocenado, que no estuvo en absoluto a la altura del interesante momento que vivió la actividad médica y científica valenciana tras la promulgación del plan Blasco.

El desarrollo que la medicina valenciana hubiera podido tener en la línea abierta por el plan Blasco quedó frustrado ante la crisis general de la España ilustrada, que culminó en los años de la guerra de Independencia y del reinado de Fernando VII. Juanto al adverso condicionamiento socioeconómico y político, conviene destacar que las reformas encaminadas a imponer un modelo universitario centralizado y uniforme perjudicaron muy gravemente a la escuela médica valenciana. En el caso de las ciencias morfológicas, la ausencia de implantación efectiva del propio plan Blasco pesó, además, negativamente. Los sucesores de Juan Bautista Poeta en la cátedra de anatomía fueron figuras tan anodinas como éste. Desde 1801 a 1807 la ocupó Vicente Soriano, quien ya había sido en los años noventa sustituto de la llamada "cátedra temporal" (o secundaria) de la disciplina, creada por el plan Blasco; desde 1808 a 1816 lo hizo Jaime Albiol, titular de dicha "cátedra temporal" de 1792 a 1796. La falta de interés de Soriano hacia las cuestiones morfológicas se refleja en el hecho de que en sus *Theses medicae* (1798) expusiera sus puntos de vista en torno a la química, la botánica, la patología y la clínica, mientras que en lo tocante a la anatomía se limitara a unas líneas en defensa del *Compendio* de Heister. Tampoco Albiol dio señal alguna de preocuparse de los avances de la asignatura de la que era profesor; hasta 1816 continuó utilizando como libro de texto dicho *Compendio*, entonces ya completamente anticuado.

La gravedad de este colapso del cultivo de la anatomía en Valencia resulta clara si se tiene en cuenta el panorama español y europeo de la época. Los últimos años del siglo XVIII correspondieron a un momento de auténtico esplendor del saber anatómico en los Colegios de Cirugía de Cádiz, Barcelona y Madrid: baste recordar, como hitos significativos, la descripción por Antonio de Gimbernat del ligamento crural que hoy lleva universalmente su nombre (1788), la publicación de los cinco volúmenes del gran tratado anatómico de Jaime Bonells e Ignacio Lacaba (1796-1800) y la del espléndido atlas osteológico del propio Lacaba, con dibujos de Isidoro de Isaura grabados por Juan Barcelón (1799). Por otra parte, durante el primer tercio del siglo XIX se produjo una profunda renovación de las ciencias morfológicas, constituyéndose como disciplinas la anatomía comparada, la nueva embriología epigenética y la anatomía textural, primero basada en la noción de tejido y luego en la teoría celular; como fechas orientadoras conviene asimismo anotar los años en los que se publicaron las *Leçons d'anatomie comparée*, de Georges Cuvier (1880-1805), *Über Entwicklungsgeschichte der Tiere*, de Karl Ernst von Baer (1828-1837), la *Anatomie générale*, de Bichat (1801), y las *Mikroskopische Untersunchungen*, de Theodor Schwann (1839). Puede afirmarse que, en este período, Valencia no solamente perdió la posición destacada que había tenido desde el Renacimiento dentro del panorama anatómico español, sino que, por vez primera desde entonces, quedó desconectada de las corrientes centrales del saber morfológico europeo. Dicho proceso condicionó precisamente la frustración del magisterio de Tomás Villanova Entraigues en la Universidad de Valencia.

Durante la guerra de Independencia, en los años que Valencia estuvo ocupada por el ejército francés (1811-1813), la enseñanza universitaria con-



*La ornitología constituyó el núcleo de la obra científica de Tomás Villanova Entraigues. Seis ejemplos de las aves existentes en las colecciones del Museo de Ciencias Naturales, de Madrid, institución en la que desarrolló su actividad docente. Son grabados procedentes de un atlas que Juan Bautista Bru había publicado en el período anterior al nombramiento del naturalista valenciano como profesor de dicho centro (*Colección de láminas que representan los animales y monstruos del Real Gabinete de Historia Natural de Madrid…*, 2 vols., Madrid, Andrés de Sotos, 1784-1786).*

tinuó funcionando, aunque con sólo 23 estudiantes de medicina en el curso académico 1811-12, cifra que ascendió a 102 el curso siguiente. La administración bajo el mando del mariscal Suchet llegó a dotar una cátedra de veterinaria y a proyectar una reforma de la profesión médica, así como a nombrar varios profesores. Entre ellos se encontraba Villanova, quien fue designado segundo catedrático de anatomía, ya que Albiol siguió al frente de la enseñanza de la asignatura como primer catedrático.

La personalidad científica de Villanova era muy distinta a la de Albiol. Hijo de Tomás Villanova Muñoz y Poyanos, catedrático de química y botánica, había nacido en Valencia en 1769. Estudió medicina en su Universidad y, tras graduarse en 1792, ganó por oposición la plaza de disector anatómico. Según testimonio de Justo Pastor Fuster "con este motivo, presentó aquel día en el teatro anatómico, como muestra de su aplicación, un pequeño gabinete de historia natural zoológica, trabajado con su mano, con la correspondiente nomenclatura según el sistema de Linneo y, además de esto, varios ojos de vidrio y en particular uno compuesto de diferentes piezas que representaban las túnicas y humores del cuerpo humano... puso también de manifiesto diferentes piezas anatómicas de muscultura y vísceras inyectadas, y otras conservadas en espíritu de vino, y algunas imitadas de cera, en especial un corazón con todas sus ramificaciones de arterias y venas". Permaneció en este puesto dos años, durante los cuales "labró varias piezas anatómicas de todo género para que sirviesen de modelo a los estudiantes", llamando en especial la atención "una estatua anatómica de cera de magnitud natural en que... se manifestaban las tres cavidades del cuerpo humano con la mayor propiedad, todas las vísceras más principales anatomizadas, las ramificaciones de las arterias y de las venas, el sistema nervioso, los vasos linfáticos y lácteos". Su competencia científica y su extraordinaria habilidad técnica explica que en 1798 fuera nombrado comisionado del Gabinete de Historia Natural del Gran Ducado de Padua, donde permaneció hasta finales de 1802 a cargo de sus colecciones y como profesor de zoología. Poco después estuvo también en relación con el Real Gabinete de Historia Natural, de Copenhague, para el que preparó en 1804 una colección de aves. Durante estos años anteriores a la guerra de Independencia no había perdido su conexión con los temas y las instituciones valencianas, como lo demuestra el hecho de que la Real Sociedad Económica de Amigos del País lo nombrara en 1807 "socio de mérito por las ciencias naturales" y que publicara, también entonces, el *Indice de las aves que forman la colección completa de las especies propias de la Albufera, disecadas y clasificadas por D. Tomás Villanova*.

El nombramiento de Villanova como catedrático de anatomía hubiera podido significar un profundo cambio en la actividad valenciana relacionada con las ciencias morfológicas. Sin embargo, como todos los profesores "intrusos" o afrancesados, fue destituido en cuanto los franceses abandonaron la ciudad, en el claustro celebrado a comienzos de septiembre de 1813. En consecuencia, su importante obra de anatomista comparado no se desarrolló en Valencia, sino en el Museo de Ciencias Naturales, de Madrid, institución en la que otro valenciano, Juan Bautista Bru, había realizado poco antes una importante obra, en especial en el campo de la paleontología y la iconografía zoológica. Daremos noticia, aunque sea resumida, de la labor que allí efectuó Villanova.

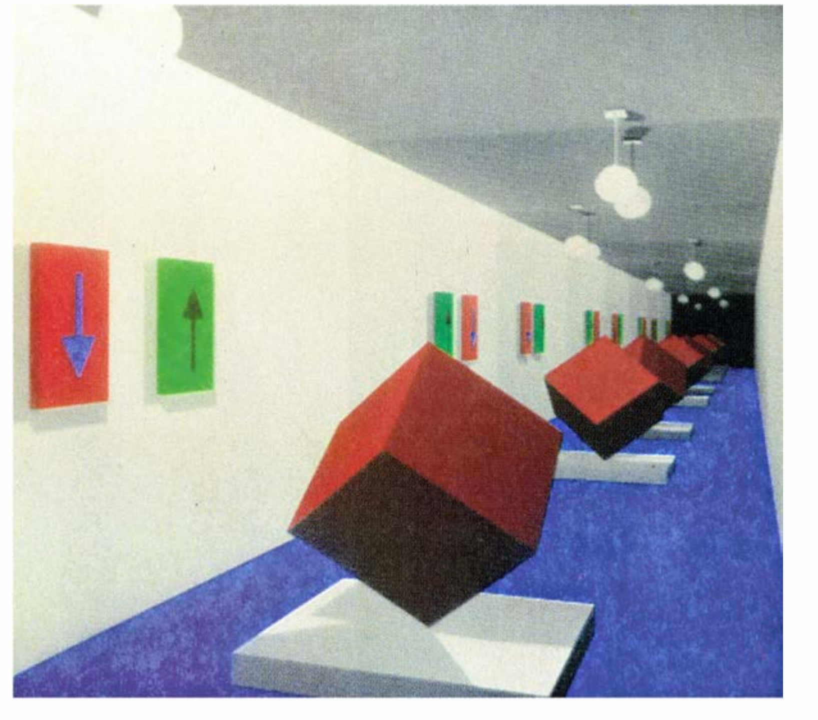
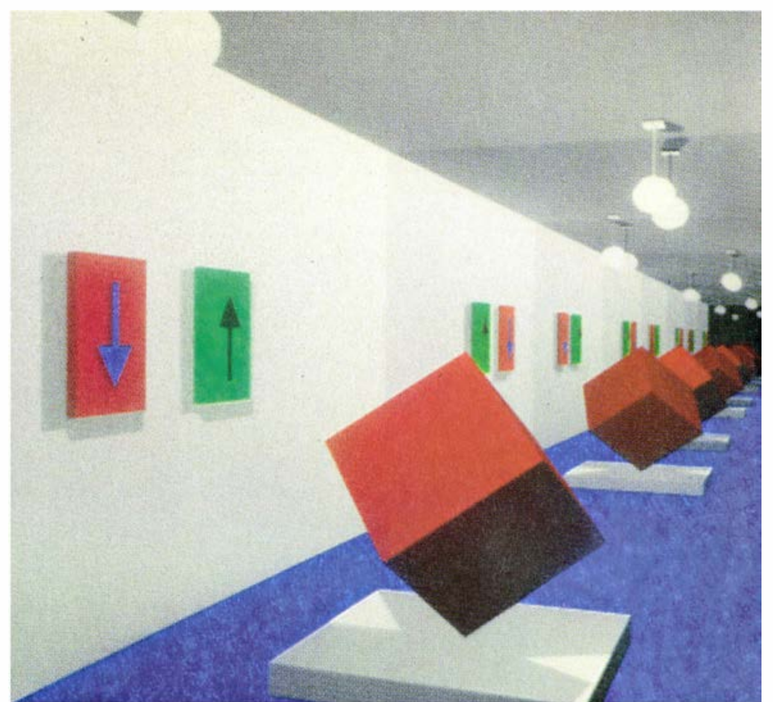
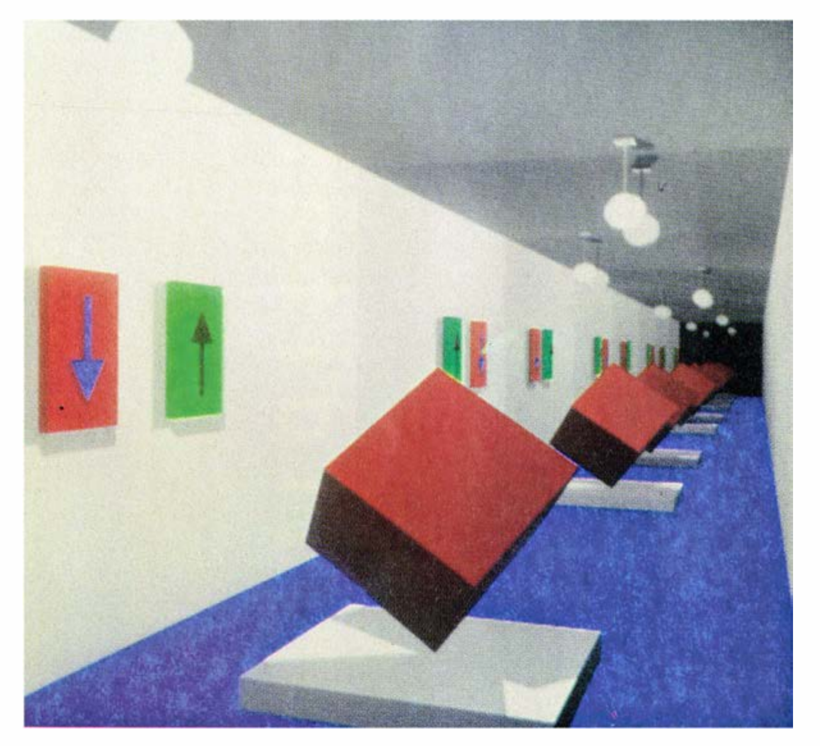
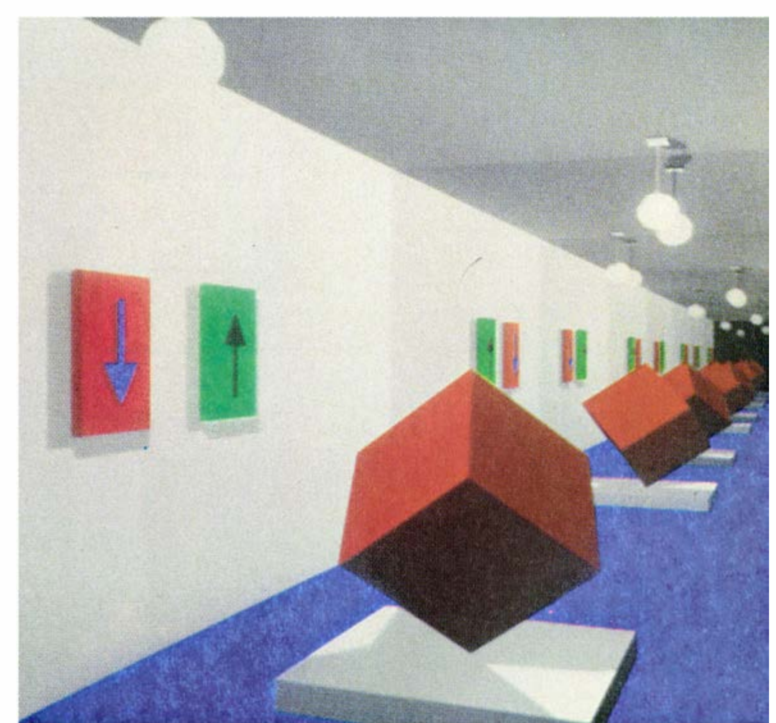
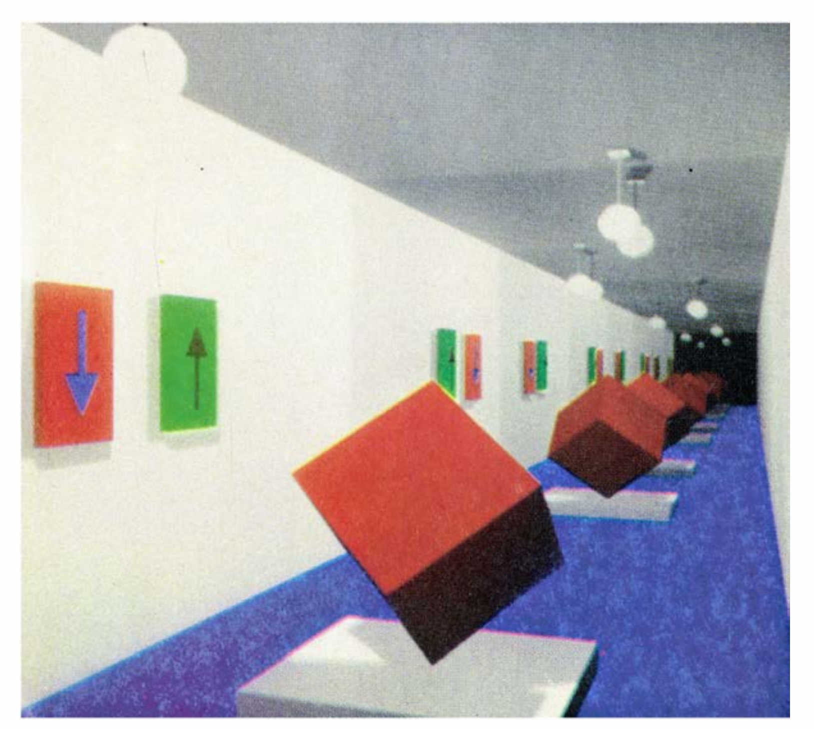
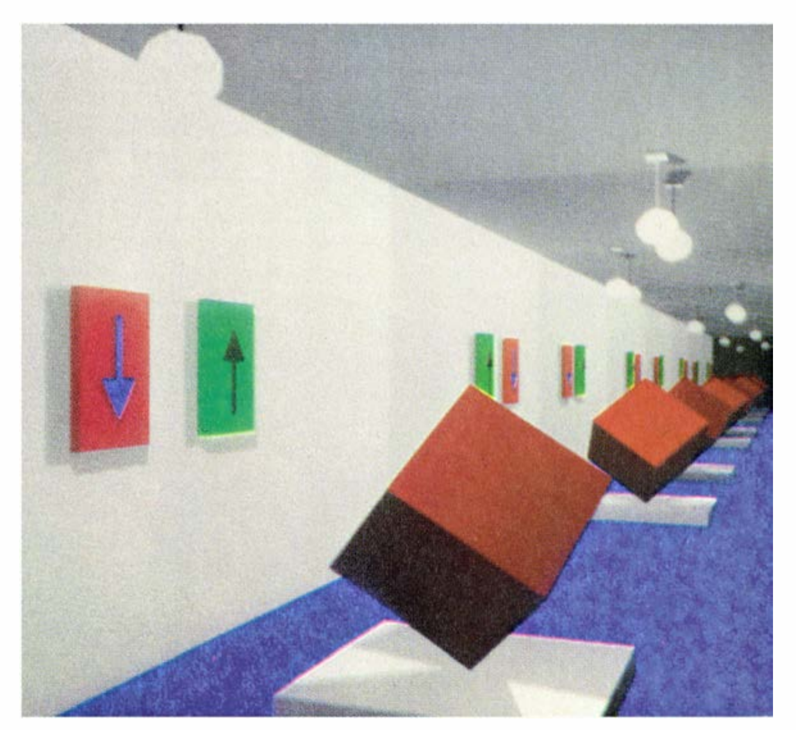
En 1819, como "viceprofesor de zoología" del Real Gabinete, Villanova dio el primer curso de anatomía comparada que se impartió en España. De acuerdo con el real decreto que lo instituyó, el método que debía seguir era el siguiente: "Un breve discurso sobre la Zoología y su utilidad servirán de apertura a sus lecciones... En seguida explicará teórica y prácticamente los conocimientos más necesarios de la anatomía comparada, para poder fundar los caracteres de la más moderna clasificación y últimamente demostrará el sistema zoológico, hasta determinar las especies en particular, todo según los últimos descubrimientos de Cuvier." Tres años más tarde, en octubre de 1822, fue nombrado catedrático en propiedad y le designaron dos "ayudantes disectores". A finales de aquel curso académico informó acerca de su docencia en estos términos: "La Anatomía Comparada ocupó mis primeras lecciones y su demostración se hizo por medio de disecciones prácticas. En seguida explané el sistema de Cuvier, presentando los caracteres de los animales sobre objetos embalsamados y existentes en el Gabinete. El sistema de Linneo fue anotado y su texto traducido literalmente a la par que se describían los objetos, y con este motivo expliqué y demostré sobre los mismos seres hasta la significación de las voces técnicas que usaba su autor. La falta de texto en el sistema del día la he suplido por medio de carpetas manuscritas que contenían los caracteres más principales y dejaba puesta delante de los objetos que determinaban, para que a su vista pudieran los discípulos copiar y repasar a un mismo tiempo por la tarde el contenido de la lección." Hasta su fallecimiento, en agosto de 1837, Villanova continuó trabajando infatigablemente en el Museo, a pesar de las desfavorables circunstancias sociopolíticas. Aparte de su dedicación a la enseñanza —que sólo interrumpió en 1832 con motivo de una grave enfermedad— consiguió enriquecer extraordinariamente las colecciones de zoología y anatomía comparada, y fue autor de una notable obra científica, aunque la mayor parte de sus trabajos no llegaron a imprimirse. El más importante fue una *Ornitología*, en tres volúmenes más un atlas, que terminó en 1825 tras dos décadas de investigaciones sobre el tema y en la que describió cerca de un millar de especies de aves.

FE DE ERRATAS

Sustitúyanse en el número 134, correspondiente al mes de noviembre de 1987, el epígrafe 6 de la página 91 y el epígrafe 9 de la página 92 por los dos siguientes:

6. EL TRANSPORTE ELECTRONICO RESPIRATORIO en los microorganismos genera una fuerza protonmotriz. En los fijadores de nitrógeno aerobios, el gradiente electroquímico de protones hace que la transferencia electrónica desde los nucleótidos de piridina hasta la Fe proteína oxidada sea termodinámicamente factible. El proceso está catalizado por la NAD(P)H: flavodoxina oxidorreductasa (*1*), una flavoproteína de la membrana cuyo grupo prostético es el FAD, y la Flavodoxina tipo II (*2*), que, una vez reducida, es el donador de electrones fisiológico de la nitrogenasa en microorganismos como *Azotobacter*.

9. EN *KLEBSIELLA PNEUMONIAE* y en otros organismos, los genes *ntr* regulan la transcripción de los genes *nif*. Un factor sigma codificado por *ntr A* se une a la ARN polimerasa y determina la transcripción de los operones activados por *C* y *A*. A su vez, *B* y *L* inhiben estas activaciones de la transcripción de los genes *nif*, previa detección de formas de nitrógeno combinado como la glutamina, siendo *L* mucho más sensible que *B* a la presencia de éstas. En presencia de altas concentraciones de amonio, *C* es producido en cantidades insuficientes para activar el operón *nif LA*, mientras que, en presencia de concentraciones más pequeñas de amonio, *L* impide la activación de los genes *nif* por *A*. La represión por O_2 está mediada por el gen *L*, pero debido a un cambio en la conformación del ADN en el que intervienen la ADN girasa y la ADN topoisomerasa I.

La próxima revolución informática

En menos de 50 años los ordenadores se han convertido en elemento fundamental de la sociedad industrial; en la etapa que se avecina crecerán al menos en un orden de magnitud y constituirán una herramienta ubicua

Abraham Peled

Cuando se inventó el ordenador electrónico, se le recibió como mera curiosidad de laboratorio. Hoy resulta inconcebible que la sociedad industrial contemporánea existiera sin él. El aparato financiero nacional e internacional, la industria manufacturera o la red de transporte, por citar algunos ejemplos, dependen de flujos electrónicos de información. Igualmente dependen del ordenador los especialistas que proyectan nuevos materiales, los farmacólogos que tratan de aislar nuevos principios activos de origen biológico o los físicos que exploran la naturaleza de la materia. La informática ha calado en el tejido de la actividad intelectual y económica en una medida sin precedentes. Posiblemente sea la informática la más apasionante aventura tecnológica de la historia. Ciertamente, así nos lo parece a quienes tomamos parte en ella. ¡Imagine el lector lo que supone trabajar en un campo que crece a ritmo exponencial!

La informática se halla ante el umbral de una nueva etapa. A lo largo del próximo decenio y merced a los progresos que registren los equipos de cómputo y la programación, incrementará, en un orden de magnitud, su potencia, versatilidad y refinamiento. Se convertirá, al mismo tiempo, en una suerte de herramienta intelectual, siempre a mano y, en última instancia, tan ubicua como el teléfono. La utilización de las máquinas se beneficiará del empleo de interfases visuales o naturales, y una red de enlaces de alta capacidad permitirá poner en contacto cualquier combinación de individuos necesitados de recursos informáticos, trátese de médicos ante un diagnóstico difícil, banqueros que redacten un contrato, ingenieros aeronáuticos en pleno diseño de una nueva célula de avión, cosmólogos interesados en un modelo de evolución del universo o estudiantes que preparan un examen.

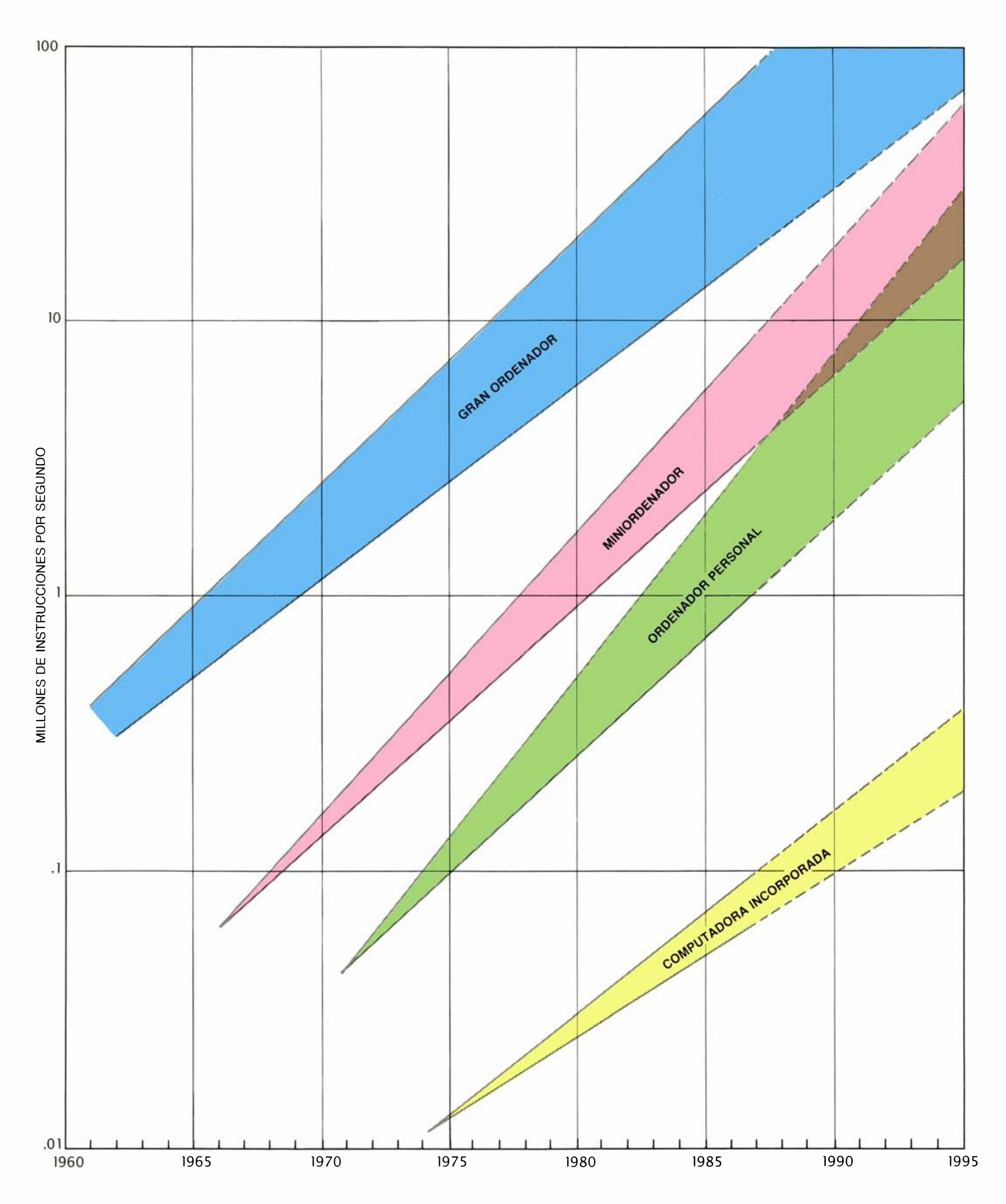
Si bien la aparición de ese útil intelectual impondrá un profundo cambio en la relación de la sociedad con el ordenador, será resultado directo de tendencias firmemente establecidas que han impulsado el progreso de la industria electrónica desde finales de la segunda guerra mundial. La más importante de esas fuerzas es la miniaturización de los componentes electrónicos, responsable principal de la persistente reducción del coste de la informática, que ha descendido a razón de un 20 o un 30 por ciento anual a lo largo de tres decenios seguidos. Al tiempo que mengua el tamaño ocupado por los transistores, puertas o cualquier otro elemento fotograbado en las pastillas de los microcircuitos integrados, aumenta en similar proporción la velocidad de funcionamiento y crece en progresión geométrica la densidad de componentes por unidad de superficie de la pastilla.

Es de esperar que el proceso de miniaturización mantenga su ritmo actual, por lo menos durante los 10 o 15 años venideros, impulsado por refinamientos ingenieriles, difíciles aunque viables, de las técnicas actuales [véase "Microcircuitos para computación avanzada", por James D. Meindl, en este mismo número]. La litografía de rayos X que emplea radiación sincrotrónica, junto con los nuevos materiales y los dispositivos de estructura superior, multiplicarán probablemente la densidad de empaquetamiento de componentes por un factor comprendido entre 20 y 40. Tales procesadores serán de seis a 12 veces más veloces que los actuales.

A esos avances acompañará un aumento constante del paralelismo en los sistemas de cómputo. La inmensa mayoría de nuestros ordenadores ejecutan sus operaciones de modo secuencial, esto es, a razón de un paso por vez. El procesamiento en paralelo resulta enormemente más potente, pues permite que el ordenador opere mucho más deprisa, al ejecutar simultáneamente la totalidad o buena parte de los pasos del problema a resolver o de la tarea encomendada [véase, más adelante, "Arquitecturas avanzadas de computadores", por Geoffrey C. Fox y Paul C. Messina].

La miniaturización marca el compás del progreso tecnológico, pero será la capacidad de producción de programas la que, en definitiva, determine con qué velocidad los sistemas de cómputo penetrarán en las iniciativas científicas, industriales y de servicios y hasta qué punto las transformarán. La razón de que así sea reside en la programación, que convierte al ordenador, de instrumento capaz de resolver en teoría un problema, en instrumento que lo resuelve en la práctica. No existe una tecnología de programación única; ni siquiera una que predomine. Los progresos llegarán del refinamiento de la estructura, de la mayor potencia de los lenguajes y de entornos de programación más eficaces [véase "Programas para computación avanzada", por David Gelernter, en este número].

1. UNA PANTALLA DE PRESENTACION muy avanzada nos muestra una recesión infinita de cubos giratorios (para reflejar el movimiento se presentan seis fotogramas). Las imágenes se confeccionaron mediante un programa concebido para SAGE (Systolic Array Graphics Engine), dispositivo especializado que constará de 16 millones de transistores. SAGE manipulará varios millares de polígonos a la vez y constituirá el prototipo de muchos dispositivos monopastilla encargados de la generación de gráficos de aquí a 10 años.



2. EVOLUCION DE LA INFORMATICA de uso general a lo largo de un período de 40 años, representada aquí mediante franjas de colores. Se ha seguido la evolución de cuatro categorías de máquinas: grandes ordenadores (*azul*), miniordenadores (*rojo*), ordenadores personales (*verde*) y computadoras integradas en otros dispositivos (*amarillo*). Cada franja define el intervalo de potencia de cómputo, en millones de instrucciones por segundo (MIPS), de que dispone cada prototipo en cualquier momento. Las líneas a trazos representan previsiones posteriores a 1987. En cualquier año que se considere, la potencia de los ordenadores centrales (*mainframes*) es mayor que la de los miniordenadores; éstos, a su vez, son más potentes que los ordenadores personales, y estos últimos, más que los dispositivos incorporados. El costo de la computación, por otra parte, baja con el descenso de potencia de la máquina. Hoy, por ejemplo, el costo aproximado relativo de la ejecución de un millón de instrucciones por segundo en un gran ordenador es de 100 unidades; en un miniordenador, de 40 unidades; en uno personal, de 3, y en una computadora incorporada (de haber alguna suficientemente potente) de 0,15 unidades. (El eje de ordenadas es logarítmico.)

Al hablar de programación, el término "estructura" alude a la eficacia, e incluso a la elegancia, con que se ensamblan las subrutinas que componen un programa completo y que permiten a la máquina llevar a cabo su tarea. A través del lenguaje el usuario le comunica a la máquina cómo poner en ejecución el programa, o parte de él, para efectuar una tarea. La potencia de un lenguaje de programación viene medida por el nivel de detalle con que deben especificarse los programas. Los lenguajes más potentes permiten que el usuario se limite a enunciar su problema en términos lógicos o matemáticos, con la seguridad de que el ordenador aportará los detalles oportunos.

El entorno de programación, constituido por el despliegue de medios materiales y lógicos mediante los cuales el programador transmite a la máquina sus instrucciones, ha evolucionado desde los conmutadores de pulsador hasta el teclado y el "ratón", desde los ordinogramas y las hojas de codificación hasta los textos y las representaciones gráficas de carácter interactivo. Los avances tecnológicos prometen facilitarnos medios más naturales de comunicación con los ordenadores [véase "Interfases para ordenadores avanzados", por James D. Foley, aquí también]. Gran parte del esfuerzo de programación se propone eliminar las dificultades mecánicas que entorpecen el uso de los ordenadores, de suerte que la única dificultad que subsista sea la conceptual, la planteada por la resolución del problema que tengamos entre manos.

El descenso vertiginoso del coste del cálculo, instado por los avances de los soportes físico y lógico, ha enriquecido la oferta de prototipos de distintos precios y prestaciones: el ordenador central, el microordenador, el ordenador personal y la computadora incorporada, instalada en el dispositivo sometido a su control.

Esos potentes sistemas han visto la luz gracias a la incesante miniaturización, tendencia que va a lomos de sutiles mejoras en las técnicas litográficas (el proceso mediante el cual se graban sobre la superficie de la pastilla los elementos que componen el microcircuito), la tecnología de procesos y las técnicas de fabricación. Hoy, los procesos litográficos y de producción alcanzan resoluciones comprendidas entre 1,0 y 1,5 micrometros. La litografía óptica creará, probablemente, dispositivos de sólo 0,4 micrometros; la litografía de rayos X podría llevarnos hasta los 0,1 micrometros. (El espesor de un cabello humano es de 100 micrometros.)

Al disminuir el tamaño de los dispositivos, hacen aparición los denominados efectos parásitos. En la escala del micrometro, los campos magnéticos y eléctricos generados por los dispositivos y por los minúsculos hilos de conexión se interfieren, hasta impedir en ocasiones el funcionamiento deseado. Ello no obstante, la tecnología de transistores de efecto de campo (crecientemente utilizada en todas las máquinas, salvo las de máximas prestaciones) seguirá alcanzando mayor densidad y velocidad. El tiempo transcurrido entre la recepción de una señal y el cambio de la salida, conocido por retardo de puerta, será del orden de los 200 picosegundos (200 billónesimas de segundo) hacia 1997; por esas fechas, la densidad será del orden de los 16 a 20 millones de dispositivos por pastilla. Consecuencia natural de tal progreso será la obtención de pastillas de memoria capaces de almacenar 16, 64 y quién sabe si 256 millones de bits cada una.

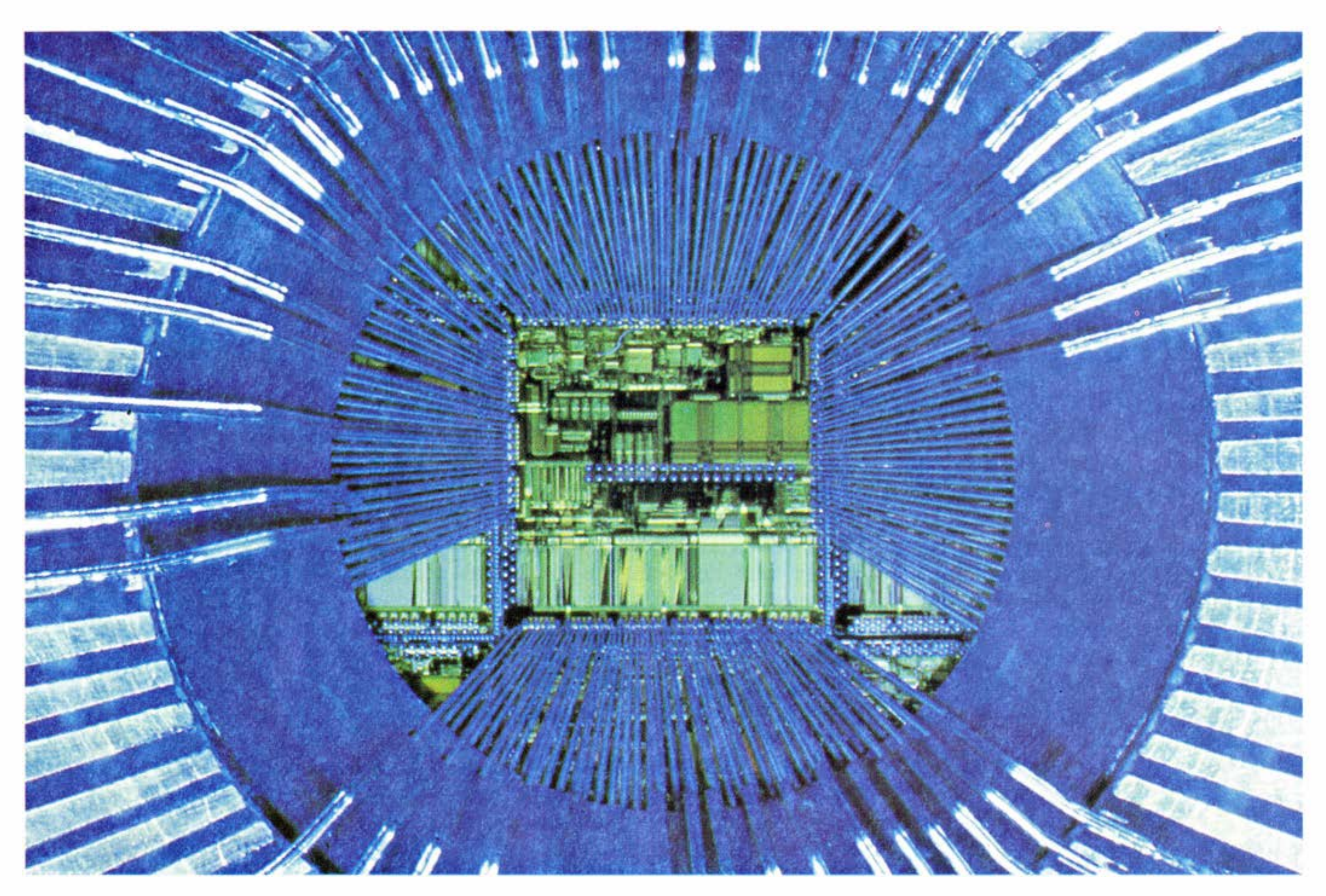
Los microprocesadores diseñados en el marco de esas tecnologías podrían operar a velocidades comprendidas entre los 30 y los 60 millones de instrucciones por segundo (MIPS). Podría fabricarse un ordenador completo en una sola pastilla, incluida la memoria y los dispositivos de entrada y salida, y capaz de funcionar a 1 o 2 MIPS: el equivalente a uno de nuestros ordenadores personales. Por supuesto, resultará factible cualquier diseño situado a medio camino entre ambos.

También podrá evolucionar en esa misma dirección la tecnología bipolar, la tecnología primigenia de los semiconductores, que subyace a los componentes principales de las máquinas de mayores prestaciones. Son previsibles puertas lógicas con demoras de 40 picosegundos. Dado que los dispositivos bipolares consumen cantidades apreciables de energía, los problemas de disipación térmica imponen límites a la densidad de empaquetamiento. Tal vez se resuelva el problema con nuevos materiales, quizás incluso con superconductores a altas temperaturas.

Al acelerarse la velocidad con que los ordenadores ejecutan las instrucciones, la máquina puede utilizar y producir mayor cantidad de información en el mismo tiempo. Ello es válido tanto si la máquina está desarrollando un modelo econométrico como si se halla generando imágenes de una molécula proteica que vibra. Es forzoso, en consecuencia, que haya que aumentar la capacidad de almacenamiento de aquellos dispositivos que han de entregar la información a los ordenadores o la hayan de recibir de éstos, una vez procesada [véase "Almacenamiento de datos para computación avanzada", por Mark H. Kryder, en este número]. En efecto, los dispositivos de registro magnético, por regla general discos, han conocido grandes mejoras en su capacidad de almacenamiento durante los últimos 20 años. En 1967, podían alojar hasta unos 30.000 bits por centímetro cuadrado; en 1987 se han alcanzado los 3.000.000. De continuar la tendencia, podrían alcanzarse densidades de almacenamiento del orden de 18 millones de bits por centímetro cúbico. Los medios de almacenamiento óptico, por ejemplo los discos compactos, proporcionan de cinco a siete veces la densidad de información de los dispositivos magnéticos.

La acelerada velocidad con la que se genera la información y el incremento de su volumen aumentan la demanda de capacidad de comunicación [véase "Avances en redes de ordenadores", por Robert E. Kahn, que aparece también en este número]. La necesidad de interconectar ordenadores y usuarios situados en diferentes puntos refuerza esa tendencia. Para afrontar la demanda esperada se está recurriendo a la tecnología de fibras ópticas. Se han tendido ya en los Estados Unidos más de 400.000 kilómetros de esas fibras. De momento sólo se utiliza un 25 por ciento de su anchura de banda, de su capacidad de transmisión, por así decirlo, principalmente para acarrear la voz. Ya han salido al mercado dispositivos capaces de introducir y extraer información de los cables de fibras ópticas a razón de 1700 millones de bits por segundo; las versiones experimentales operan a velocidades de 10.000 millones de bits por segundo.

Los circuitos de optoelectrónica integrada basados en el arseniuro de galio (un material semiconductor capaz de transformar eficazmente la luz en corriente eléctrica, y viceversa) podrían ofrecer importantes mejoras de velocidad y costo. En el próximo decenio se dispondrá ampliamente de redes informáticas de alta velocidad, con capacidades de 45 millones de bits por segundo. Los efectos de la puesta en servicio de esos sistemas, tres órdenes de magnitud más potentes que sus predecesores de la década de 1970, no se limitarán a los meros cambios de es-



3. TODO UN ORDENADOR, integrado en una sola pastilla. Mide 10 milímetros de lado y puede operar a 350.000 instrucciones por segundo. El dispositivo es uno de los 87 grabados sobre la misma oblea de semiconductor. Los conectores y las sondas permiten a los técnicos de la División de Investigación de IBM ensayar simultáneamente todos los circuitos del dispositivo. Si alguno aparece defectuoso, se desecha al cortar la oblea para separar los ordenadores.

cala. Habrá aplicaciones que comporten el movimiento rápido de imágenes visuales. Y podrán estar a punto nuevas configuraciones de equipos informáticos en las que intervengan máquinas muy alejadas entre sí, pero cuya rápida respuesta en nada envidiará a las que se encuentran en el mismo edificio. La utilización de esas redes exigirá considerables avances en los programas de gestión de comunicaciones y nuevos y más eficientes protocolos, así como elementos especializados de cómputo, que incorporen físicamente tales protocolos a los equipos informáticos.

Los progresos en el campo de los semiconductores y en la resolución de las pantallas permitirán a los usuarios de tales sistemas, y de otras instalaciones de cómputo, trabajar con pantallas rápidas, de alta calidad y de precios razonables. Las pantallas ofrecerán imágenes vívidas, que plasmarán los resultados de complejas simulaciones y modelizaciones. Podremos relacionarnos con la pantalla y explorar qué efectos provoca un cambio, así la introducción de un elemento nuevo en una estructura molecular o la modificación del alabeo de un ala. La alta resolución ofrece otra ventaja: la pantalla podrá presentar textos con la calidad de una página cuidadosamente impresa. Los estudios realizados por la División de Investigación de IBM han puesto de manifiesto que esa mejora, que aumentaría la calidad del texto entre un 20 y 30 por ciento, daría por fruto velocidades de lectura y tasas de retención comparables con las obtenidas en medios impresos.

Tales dispositivos exigirán el desarrollo de microcircuitos especiales para el tratamiento de gráficos. En este sentido, estamos desarrollando en IBM un prototipo de esa clase de sistema. Lo denominamos SAGE (por *systolic array graphics engine*). SAGE manipula un juego de varios millares de polígonos, les imprime cambios de forma y los anima. Mediante computación en paralelo genera al instante una nueva imagen en función de las necesidades. SAGE consta de más de 16 millones de transistores, que algún día se litografiarán sobre una pastilla.

"¡Ojalá progresáramos en la obtención de los programas con la rapidez con que lo hacemos en la construcción de mejores ordenadores!" Ante tales lamentaciones cabe responder que se han dado pasos importantes en la capacidad de redacción de programas, sólo que esos progresos han quedado empequeñecidos por el ritmo sin precedentes con que ha avanzado el material de soporte y desbordados por las nuevas posibilidades que éste ha creado. La confección de programas es intrínsecamente difícil, pues comporta especificar la solución de un problema con detalle suficiente para que la máquina la lleve a cabo. La tarea de elegir entre las múltiples opciones posibles, teniendo presentes al mismo tiempo restricciones tales como la forma en que se van a recibir los datos, constituye un duro esfuerzo teórico. La complejidad de la tarea, sumada al hecho de que los programas de resultados satisfactorios se modifican continuamente para actualizarlos y adaptarlos a equipos nuevos, dificulta aún más la cuestión.

La creación y evolución de lenguajes de alto nivel forman parte de la solución al problema que plantea lograr que los programas extraigan del material todo el potencial posible. Los lenguajes de alto nivel expresan sucintamente las instrucciones y proporcionan, en comparación, pocos detalles.

Por ejemplo, en vez de indicarle específicamente al ordenador que revise cada entidad, se le podría comunicar, en el lenguaje SETL, que realizase una operación matemática con un grupo de valores: "Calcula la cuota fiscal en porcentaje de los ingresos brutos correspondientes a las declaraciones cuyos ingresos brutos sean de 10.000.000 de pesetas o superiores". Tales lenguajes son populares en el diseño y construcción de prototipos iniciales de sistemas. La construcción de prototipos iniciales contribuye a identificar y resolver muchas dificultades conceptuales. Se trata de un procedimiento, muy prometedor, de ir construyendo programas gradualmente, que podría mejorar bastante la productividad.

Otra razón de mejora vendrá de la inclusión de lenguajes encaminados a facilitar el desarrollo de grandes proyectos de programación. La dificultad estriba en que, en tales proyectos, por tiempo y economía, muchas partes del sistema tienen que diseñarse en paralelo. La restricción obliga a centrar la atención en un diseño cuidadoso de las interfases. Considerada la longevidad de esos programas, el lenguaje tiene que dar acomodo a sus modificaciones y evoluciones. En este campo, casi todo el acento se carga en asegurar el porvenir de los métodos de encapsulamiento de datos y en establecer procedimientos que permitan que los datos sólo se utilicen de ciertas formas. Con frecuencia resulta necesario llegar a una suerte de compromiso entre la importancia de tales consideraciones y la necesidad inmediata de un funcionamiento adecuado.

En el pasado, las decisiones se inclinaron del lado de la eficacia de funcionamiento. Mas, al ir abaratándose el costo de la computación, las decisiones favorecerán la estructuración y modularidad de los programas. Finalmente, se asistirá al desarrollo y explotación de métodos de programación que reduzcan el tiempo requerido para instruir al ordenador. Vale de ejemplo de ello el lenguaje PROLOG, que libera al programador de los detalles de secuenciación y le permite concentrarse en las relaciones lógicas sencillas entre elementos del programa.

La batalla del programador contra esos escollos irá simplificándose merced a la potencia combinada de los ordenadores personales y los bancos de información de fácil acceso que suministren información acerca de los programas. Tales sistemas, sumamente interactivos, proporcionarán rápido acceso a múltiples aspectos del programa. Se advierte de inmediato ese carácter interactivo en el sintetizador de programas desarrollado por la Universidad de Cornell, donde trabajos previos pusieron de manifiesto la eficacia de un conjunto de herramientas de índole lingüística en la adquisición de destreza programadora. Más recientemente, el Garden Project, de la Universidad Brown, explota la capacidad gráfica de las terminales de trabajo para presentar al usuario múltiples vistas del programa.

Los sistemas expertos constituyen, a la vez, una herramienta y un producto

LA ECONOMIA DEL DISEÑO DE SISTEMAS

Para adquirir un sistema capaz de realizar una tarea a costo mínimo, el jefe de sistemas informáticos mencionado en el texto podría valerse de las siguientes ecuaciones.

(1) $P_1 = M \times C(M)$ siendo

P_1 = Precio de la ejecución de la tarea con un solo procesador
M = Megainstrucciones por segundo (MIPS) necesarias para ejecutar la tarea
C = Precio unitario en $/MIPS correspondientes al valor pico de las MIPS especificadas

(2) $P_N = (e \times M) \times C\left(e \times \frac{M}{N}\right)$ siendo

P_N = Precio de ejecución de la tarea en varios procesadores más lentos
N = Número de procesadores más lentos
e = Factor de expansión, que es el aumento de trabajo requerido a consecuencia de dividir la tarea entre varios procesadores

La opción multiprocesador es preferible si:

(3) $e < \frac{C(M)}{C(e \times M / N)}$

Con $N = 10$ y $e = 2$, la opción multiprocesador es preferible si:

(4) $\frac{C(M)}{C(M / 5)} > 2$

Una descomposición más completa de los gastos del banco sería:

(5) $P_N = (k \times M + m \times I_m) \times C\left(\frac{k \times M + m \times I_m}{N}\right) + m \times B$ donde

k = Expansión de MIPS correspondiente a la aplicación
m = Número de mensajes a intercambiar entre los procesadores
I_m = Millones de instrucciones necesarias por mensaje promedio
B = Precio de la comunicación por mensaje promedio

Con $N = 10$, $m = 100$, $k = 2$, $I_m / M = 0{,}005$ y $m \times B/C(M) = 0{,}1$, la opción multiprocesador es superior si:

(6) $\frac{C(M)}{C(M / 4)} > 2{,}8$

Si B aumenta cinco veces, las condiciones cambian a:

(7) $\frac{C(M)}{C(M / 4)} > 5$

Si I_m se reduce 10 veces, las condiciones cambian a:

(8) $\frac{C(M)}{C(M / 4{,}87)} > 2{,}27$

Si $m \times B/C(M) \geq 1$, *jamás* es preferible dividir la tarea entre varios procesadores.

de la tecnología de la programación. Quizás el campo donde más amplio uso se haga de esos sistemas sea el de la creación de programas. Un sistema experto típico consta del planteamiento lógico de la aplicación (a saber, un conjunto de reglas del tipo "si se dan *A* y *B*, ejecútese *C*") más un programa de uso general (la máquina de inferencias que invoca las reglas apropiadas para alcanzar un objetivo especificado y seguidamente las ejecuta). Este estilo de redacción libera al programador de la necesidad de detallar las estructuras de los datos y de establecer el flujo del control (la especificación detallada del orden en que deben ejecutarse las instrucciones). Se puede también introducir reglas en cualquier momento y en cualquier orden. Estas características hacen que los sistemas expertos resulten idóneos para la elaboración rápida de prototipos y la construcción gradual de la aplicación. El intenso trabajo que se está desarrollando sobre sistemas expertos en el campo de la medicina está acercando la tecnología a su empleo práctico [véase "Computación avanzada en el campo de la medicina", por Glenn D. Rennels y Edward H. Shortliffe, en este número].

Aunque el soporte lógico, la programación, ha ido siempre a la zaga del sopote físico, una mirada retrospectiva a la evolución de los ordenadores nos muestra que uno y otro han producido un repertorio creciente de capacidad computacional; la tendencia revela, asimismo, todos los signos de un crecimiento vigoroso [*véase la figura 2*]. Se ha registrado un crecimiento constante de la potencia de cómputo y se ha incrementado paulatinamente la magnitud de la capacidad ofrecida en los principales márgenes de precios. La tecnología ha permitido la construcción de ordenadores bastante menores, pero "útiles", a un precio notablemente inferior por millón de instrucciones por segundo, así como a un precio global del sistema inferior también. Han surgido, pues, nuevas categorías de precios: al ordenador de tamaño imponente siguieron, primero, los miniordenadores y, después, los ordenadores personales y los dispositivos con procesador incorporado. La configuración de esas máquinas de cómputo es fruto natural de la tecnología y la economía. La mayor densidad de los circuitos y su menor consumo de energía, sumados a un encapsulamiento barato y una simplificación de la refrigeración, han posibilitado nuevos objetivos de diseño. La rebaja de los costes incrementa la demanda de capacidad de cómputo, generándose así las economías de escala que derivan de los grandes volúmenes de producción.

Importa recordar que los sistemas pequeños poseen, proporcionalmente, menos memoria, menos capacidad de almacenamiento en disco y menos capacidad también de entrada y salida. Así pues, no pueden ejecutar muchos programas y aplicaciones grandes. Por el contrario, los usuarios cuyas tareas resultan ejecutables en esos sistemas pueden elegir entre un número creciente de opciones; a igualdad de los res-



4. RP3, banco de pruebas de ordenadores en paralelo construido por IBM y la Universidad de Nueva York, que simula diversas arquitecturas en paralelo y las enfrenta con tareas específicas. Así se determina la eficacia de una arquitectura para una aplicación y se evita el gasto de la producción de un prototipo.

tantes factores, las nuevas aplicaciones tienden a aparecer en la más baja de las categorías viables. Conforme aumenta el número máximo de MIPS, crece con mayor celeridad el número total de aplicaciones económicamente justificadas: es éste la suma de todas las aplicaciones justificadas en cada nivel.

La experiencia recabada hasta la fecha manifiesta que los sistemas más versátiles son los provistos de un único procesador (o de un pequeño número de procesadores estrechamente acoplados); es más, su capacidad de almacenamiento, su capacidad de procesamiento y su tiempo de respuesta son predecibles y se comportan bien en una amplia gama de aplicaciones. El rendimiento, capacidad de procesamiento y tiempo de respuesta de cualquier configuración con varios procesadores dependen estrechamente de la aplicación que se les asigne.

Supóngase, por ejemplo, que un director de sistemas de información bancaria desea adquirir un sistema con capacidad para operar a 100 transacciones por segundo, al objeto de controlar las retiradas de fondos de una red de cajeros automáticos. Podría optar por una sola máquina de 50 MIPS para lograr un tiempo de respuesta inferior al segundo. Es muy probable que 10 máquinas de 5 MIPS resulten más económicas, pero a su precio el directivo debe sumarle el factor de expansión, esto es, el costo de la programación y del equipo necesario para coordinar la actividad de los 10 ordenadores. Al tomar en consideración el factor de expansión, pudiera muy bien resultar que la máquina de 50 MIPS fuera más económica [*véase el recuadro de la página 17*].

Puede ocurrir también que el costo de la comunicación entre los ordenadores sea tan elevado que iguale o supere el costo de un ordenador de capacidad igual a la total de las distintas máquinas. La cuestión es que, conforme crece el número de formas de sistemas informáticos al alcance, la elección depende en gran medida de las pautas de acceso a los datos y de compartición de los mismos, y no tanto del costo bruto de la computación.

Consideraciones como las anteriores pesan en las decisiones relativas a sistemas tan dispares como son los que respaldan las operaciones de una empresa o los que sirven para reforzar la creatividad del usuario privado. El sistema de una empresa puede constar de más de 100.000 nodos, capaces de generar 1000 peticiones simultáneas, cifras que se incrementan en un orden de magnitud si se manejan además transacciones entre empresas distintas.

La elección de sistema dependerá estrechamente de las características de la aplicación. Un sistema para la ejecución de transacciones comerciales puede exigir tiempos de respuesta de un segundo. Pudiera así suceder que, para determinado volumen de transacciones y un precio dado, el sistema más económico constara de cuatro procesadores, capaces de ejecutar cada uno 18,5 MIPS. La distribución geográfica de un tal sistema impone siempre un gravamen. El tiempo de respuesta sube en proporción al número de mensajes asociados a cada transacción. Incluso dando por supuesto que las líneas de comunicación posean tal ancho de banda que no se registren demoras por transmisión, el consumo indirecto de computación (las instrucciones que es preciso ejecutar en cada ordenador para enviar o recibir el mensaje) que exige el intercambio de mensajes dilatará el tiempo de respuesta. Los ingentes sistemas de cobertura mundial plantearán varios retos principales: distribución de la programación, mantenimiento y puesta al día de la misma, más todos los problemas suscitados por el funcionamiento continuo y la entrada de grandes empresas en la red. La complejidad que aparece al aumentar de escala exigirá nuevos algoritmos, nuevas estructuras de datos y nuevos planteamientos teóricos.

En contraste con los sistemas capaces de atender a una empresa entera, en los cuales lo esencial es la compartición de datos, lo que más importa en los sistemas de computación de uso personal es la potencia. Tales sistemas están diseñados para ayudar a un individuo o a un pequeño grupo de profesionales en la resolución de un problema mediante la aplicación de potencia computacional. Sus raíces se encuentran en la computación interactiva y de tiempo compartido de los grandes ordenadores, pero su crecimiento ha sido explosivo desde el advenimiento del ordenador personal. Se espera de ellos cambios profundos de la naturaleza de la computación, que convertirán al ordenador en amplificador universal de la inteligencia humana.

El ordenador personal típico del próximo decenio alcanzará una potencia de cómputo quizás un orden de magnitud superior a la de la máquina típica de nuestros días, y una capacidad de almacenamiento entre 10 y 100 veces superior. Su pantalla, de alta calidad y velocidad, reproducirá en menos de un segundo una imagen formada por entre un millón y cuatro millones de elementos de imagen, pixeles, a elegir cada uno de entre 64 y 256 matices cromáticos.

Una porción considerable de la potencia de cómputo se destinará a facilitar el manejo de la máquina, dando cabida a una interacción entre hombre y máquina más natural para el operador, basada, por ejemplo, en la voz o en la escritura a mano. Los especialistas de la División de Investigación de IBM han puesto a punto un sistema capaz de reconocer 20.000 palabras si se pronuncian separándolas con una breve pausa. El sistema comprende 60 millones de instrucciones, que activan cuatro microprocesadores especializados y un ordenador IBM PC/AT. Hace cuatro años, tal ingenio hubiera ocupado una habitación; dentro de cinco, posiblemente quepa en una tarjeta. El reconocimiento continuo del habla exigiría una potencia de cómputo unas 30 veces superior, y además se precisarían una capacidad de almacenamiento y potencia adicionales para mejorar la precisión y reducir la sensibilidad al ruido del ambiente.

Los terminales que imitan el papel permiten al usuario "escribir" sobre una pantalla plana de cristal líquido. El ordenador reconoce los caracteres y los traduce a órdenes, texto o dibujos. El sistema prototipo exige de dos a cuatro MIPS; los algoritmos necesarios, siquiera mejorados, consumirán sólo una pequeña fracción de la capacidad de procesamiento de los futuros ordenadores personales.

Además de reconocer el habla y la escritura, cada vez se sacará más partido de la potencia del ordenador personal para mostrar los resultados de un cómputo de modo visualmente intuitivo. Los gráficos de sectores y los esquemas se enriquecerán con imágenes en perspectiva, presentadas en distintos matices cromáticos. Tales sistemas interactivos, muy poderosos, determinarán que el ordenador personal se utilice como punto de encuentro natural entre los humanos y los grandes ordenadores: una suerte de ventana abierta a una gran red informática o a un superordenador, capaz de mostrar en cuestión de segundos los resultados de horas de cálculos.

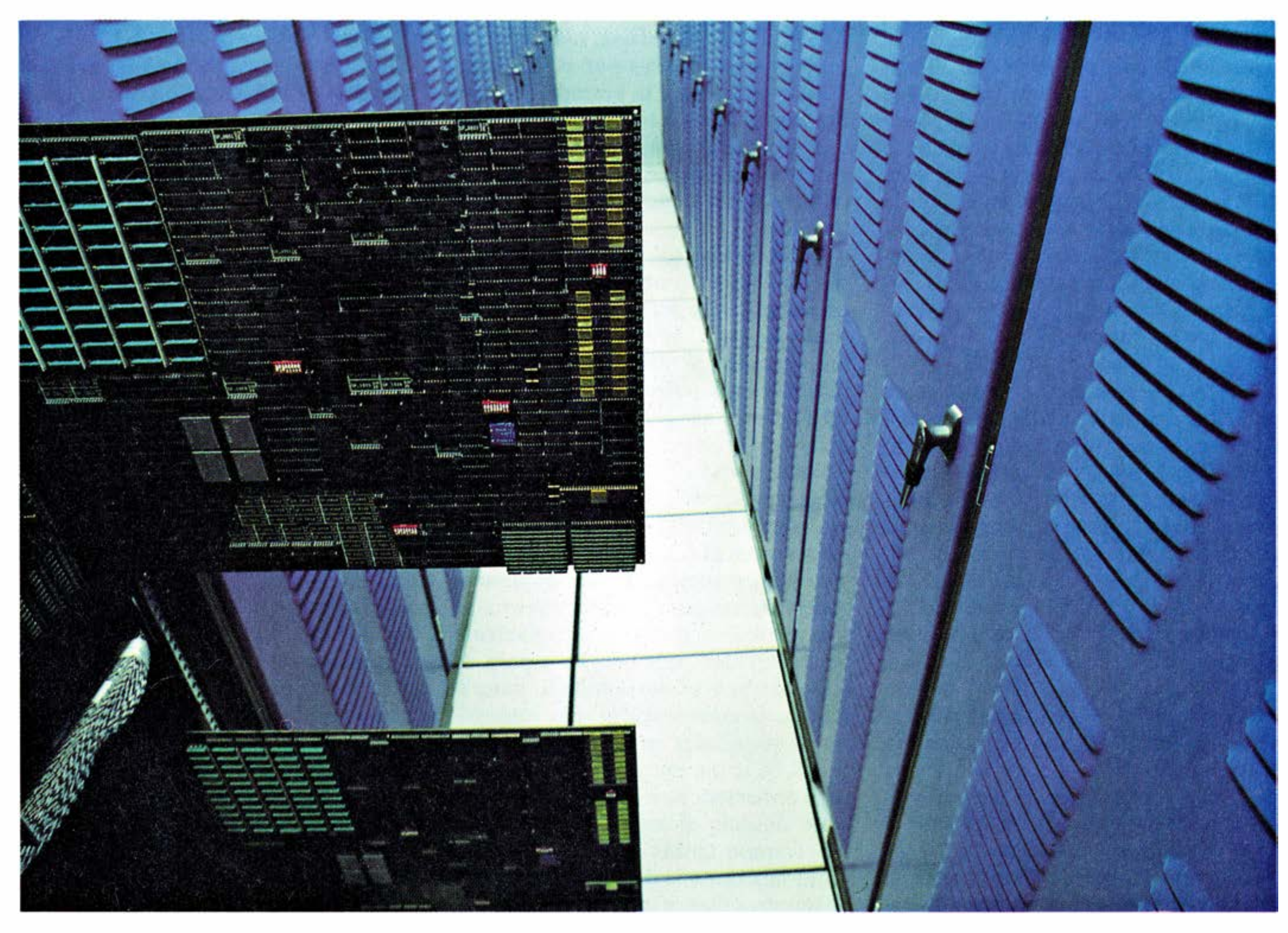
La disponibilidad de una capacidad de computación de gran potencia y flexibilidad permitirá a los ingenieros mo-

delizar fenómenos tan complejos como la conducta y el rendimiento de una estructura aeronáutica en diversas condiciones de vuelo [véase "Computación avanzada en la industria", por Albert M. Erisman y Kenneth W. Neves, en este número]. El usuario suele hacerse así una idea más clara y penetrante que a través de la mera simulación o el propio experimento. También más barata. La capacidad de modelizar electrónicamente el comportamiento y el resultado de sus proyectos prestará un innegable servicio a arquitectos, decoradores y escenógrafos. Tal modelización mediante ordenador acelerará la evolución desde la idea hasta el producto. Las nuevas ideas así adquiridas podrían llevar a productos enteramente nuevos o a nuevos conceptos de diseño. Gracias a la simulación electrónica, el investigador realizará experimentos imposibles de llevar a cabo en la práctica como, por ejemplo, determinar qué sucede tras el choque de dos galaxias [véase, más adelante, "Computación avanzada en el campo de la ciencia", por Piet Hut y Gerald Jay Sussman].

El lugar donde se efectúe el cómputo dependerá de la magnitud del mismo y del tiempo de respuesta requerido. A la hora de modificar la forma de un automóvil, el ingeniero podría apelar a 60 millones de instrucciones para calcular el resultado y a cuatro millones de instrucciones más para proyectarlo en su pantalla desde un ángulo determinado. En un ordenador personal que funcione a 1 MIPS, la tarea podría ejecutarse en 64 segundos. En un ordenador de 20 millones de ilustraciones por segundo, al cual estuviera conectado el personal por un enlace de alta velocidad, la presentación requeriría cuatro segundos, mientras que los 60 segundos exigidos por el cómputo se reducirían a tres. El compromiso se establece, pues, entre el tiempo del individuo y el costo de la computación. Con el abaratamiento de los precios de computación, los dispositivos se dedicarán a mejorar la productividad del individuo.

Otra transacción implícita: la que se da entre la modelización informática de la carrocería del automóvil o de la célula del avión y la construcción de una versión experimental para ensayo en el laboratorio. Cuando se disponga de las formidables potencias de cómputo necesarias para simular la realidad, se aplicarán a todas las fases del ciclo que median entre el diseño y la producción, a excepción, lógicamente, de las etapas finales. La experiencia demuestra que la simulación computarizada resulta mucho más flexible y económica que la construcción de prototipos. El resultado es, invariablemente, un diseño mejor de cualquier objeto, desde la quilla de un velero hasta un ordenador más rápido.

Fruto maduro del crecimiento exponencial de la densidad de empaquetamiento y de la profundización del refinamiento de la programación es el procesador en paralelo. Sin duda, el ordenador monoprocesador, el tradicional, seguirá constituyendo la máquina de cómputo mayoritaria durante algún tiempo. Los monoprocesadores son máquinas versátiles y de buen comportamiento. No obstante, las razones económicas, sumadas a los progresos de la informática, crearán muchas aplicaciones útiles para los procesadores en pa-



5. GF11, ordenador científico en paralelo que se está construyendo en IBM; realizará 11 millones de operaciones con coma flotante por segundo. Su primera tarea consistirá en deducir la masa del protón. Se espera que lo logre en cuatro meses, cuando un superordenador tardaría 15 años para esa misma tarea.

ralelo (así como para las máquinas de inteligencia distribuida).

Decididamente, se avanza en pos de las máximas prestaciones absolutas. Se están explorando cuantos esquemas de interconexión de elementos computacionales se suponen posibles, desde millares de elementos baratos y relativamente lentos hasta decenas de los más rápidos monoprocesadores construidos. Para minimizar los tiempos muertos resulta esencial establecer líneas de interconexión de banda muy ancha, que permiten que los datos y las instrucciones pasen rápidamente de un procesador a otro. Se enlazan ya algunos procesadores de muy alta velocidad mediante circuitos optoelectrónicos que excitan fibras ópticas. Cuando el número de procesadores se eleva, hay que conectarlos a través de una central de conmutación (por ejemplo, una matriz densa de transistores de efecto de campo), capacitada para enlazar dos procesadores cualesquiera, a imagen del tendido telefónico, que pone en comunicación dos receptores. Sería imposible crear una red de conexiones directas entre todos los procesadores. Como el lector sospechará, tales redes de comunicaciones son hoy objeto de intensa investigación. Un modelo sumamente interesante propone que la red de comunicaciones se organice de acuerdo con las pautas cerebrales de los mamíferos.

Las herramientas de análisis de rendimiento, bastante refinadas para el caso de los monoprocesadores, resultan todavía rudimentarias en los procesadores en paralelo. En IBM se está ensayando la interconexión de unos pocos procesadores de alta velocidad y la construcción de prototipos altamente paralelos. Uno de esos dispositivos, RP3, consiste en un procesador en paralelo cuya flexible organización le permite servir de campo de pruebas para distintas arquitecturas en paralelo. En el desarrollo de RP3 colabora la Universidad de Nueva York; el proyecto cuenta con una subvención de la Defense Advanced Research Projects Agency. La puesta en servicio del ordenador permitirá recoger información en tiempo real sobre la capacidad de distintas clases de máquinas en paralelo para enfrentarse a una amplia gama de aplicaciones.

Una máquina de otro tipo, surgida también de los laboratorios de investigación de IBM, es YSE. Se trata de un procesador en paralelo que simula los circuitos en el nivel de componente, esto es, de interruptores y puertas lógicas, y no en el nivel arquitectónico. Altamente paralelo, este procesador ejecuta una simulación centenares de veces más deprisa que cualquier monoprocesador al uso. YSE posibilita el ensayo y depuración rápidos de varios diseños sin tener que afrontar los gastos de construcción de prototipos.

La tendencia hacia la construcción de ordenadores especializados se acelerará conforme vaya disponiéndose de microprocesadores cada vez más potentes. En esa línea, IBM está construyendo la máquina Gigaflop 11, capaz de efectuar 11 gigaflops por segundo. (Flop es abreviación de *floating point operation*, es decir, operación con números representados en la notación de "coma flotante", que es la forma habitual de representar números reales los ordenadores.) La red de conmutación denominada "de permutación" se encarga de trasladar los datos entre los 576 procesadores, que ejecutan una misma operación obedeciendo a un plan predeterminado. El primer desafío que habrá de afrontar GF11 será el cálculo de la masa del protón según la cromodinámica cuántica, teoría que trata de describir la estructura última de la materia. Semejante cálculo, que exige aproximadamente 10^{17} operaciones "de coma flotante" para lograr una precisión del 10 por ciento, exigiría unos 15 años de funcionamiento continuo en un superordenador típico. GF11 debería ejecutar esa tarea en cuatro meses.

Los avances en los sistemas de cómputo persistirán (quién sabe si creciendo exponencialmente, pero desde luego sin aflojar el paso) los 10 o 15 próximos años, cuando menos. La amplia disponibilidad de ordenadores para una comunidad cada vez mayor de usuarios reforzará la creatividad de éstos y alimentará el progreso.

Aunque la informática actual sólo puede dilatar habilidades mentales sencillas y bastante rutinarias, se avanza ya con decisión hacia el reforzamiento informático de otras de carácter más analítico e inferencial. Igual que las máquinas capaces de extender y multiplicar las posibilidades físicas de los humanos facilitaron la Revolución Industrial, también la informática, por su capacidad para ampliar nuestras dotes mentales, es el motor que impulsa la revolución en curso, una revolución que todavía no ha encontrado el nombre que la describa adecuadamente. Porque el viaje apenas ha comenzado.

2587000225

102575

MADE IN U.S.A.

COPYRIGHT 1986

Arquitecturas avanzadas de computadores

Los computadores al uso resuelven los problemas ejecutando una sola operación cada vez. Los ordenadores avanzados resuelven varias partes del problema a un mismo tiempo

Geoffrey C. Fox y Paul C. Messina

Supongamos que el lector ha comprado un terreno donde ha decidido levantar una casa y ha contratado a un solo operario para realizar todos los trabajos. El obrero procedería secuencialmente: albañilería, fontanería, electricidad y pintura. Un trabajo después de otro y, cada uno, con cierto orden.

A ojos vista, se trata de una forma innecesariamente parsimoniosa de edificar una casa. Varios trabajos, los de albañilería por ejemplo, terminarían antes si se dividieran entre diversos operarios, que actuasen al mismo tiempo. Otras tareas, tales como la instalación eléctrica y la fontanería, son independientes entre sí y podrían acometerse a la vez por distintas cuadrillas. Pues bien, ese lentísimo método secuencial –un obrero ejecutando el trabajo entero, del que desarrolla una sola parte cada vez– constituye una analogía adecuada de la forma en que proceden la mayoría de nuestros ordenadores.

La mayoría posee, en efecto, un solo procesador. Constituye éste la unidad que realiza los trabajos de computación: suma, multiplicación o comparación de dos números, por ejemplo. Al programar, el hombre descompone cada tarea informática en una secuencia de pasos –un programa– y el procesador la ejecuta fase a fase. Este proceso es forzosamente lento por dos razones. En primer lugar, durante cada fase de una computación, gran parte del procesador permanece inactivo; para multiplicar dos números, por ejemplo, se requieren varios pasos, y el circuito que ha de ejecutar uno quedará inactivo mientras se realice otro paso. En segundo lugar, y volviendo a la analogía de la construcción de una casa, muchas tareas de computación pueden dividirse en subtareas independientes entre sí y susceptibles de ser acometidas, al mismo tiempo, por procesadores diferentes.

Los especialistas en arquitecturas avanzadas están abordando procesos que permiten obviar ambas fuentes de lentitud. Por un lado, están ideando formas de acelerar la velocidad del procesador manteniendo activo más de un componente. Por otro, están diseñando sistemas que ponen en conexión varios procesadores para constituir lo que se ha dado en llamar procesadores o computadores paralelos. En ambos enfoques se pretende que concurran varias fases de la computación.

¿Cómo acelerar la velocidad del procesador? En los procesadores al uso la lentitud se produce, sobre todo, en el acceso a la memoria. Si hay que recoger datos o instrucciones del banco de memoria antes de ejecutar cierta operación, todas las unidades funcionales del procesador (los circuitos lógicos responsables de los distintos pasos de computación) permanecen inactivos mientras actúa la memoria. Terminada la operación, habrá que salvar en la memoria el resultado abriéndose otro intervalo durante el cual las restantes unidades funcionales permanecen quietas. El problema se soluciona con máquinas en las que, mientras se ejecuta una operación, el siguiente conjunto de instrucciones se recoge de la memoria, se decodifica o se fragmenta en subtareas, de las que se encargan diferentes unidades funcionales.

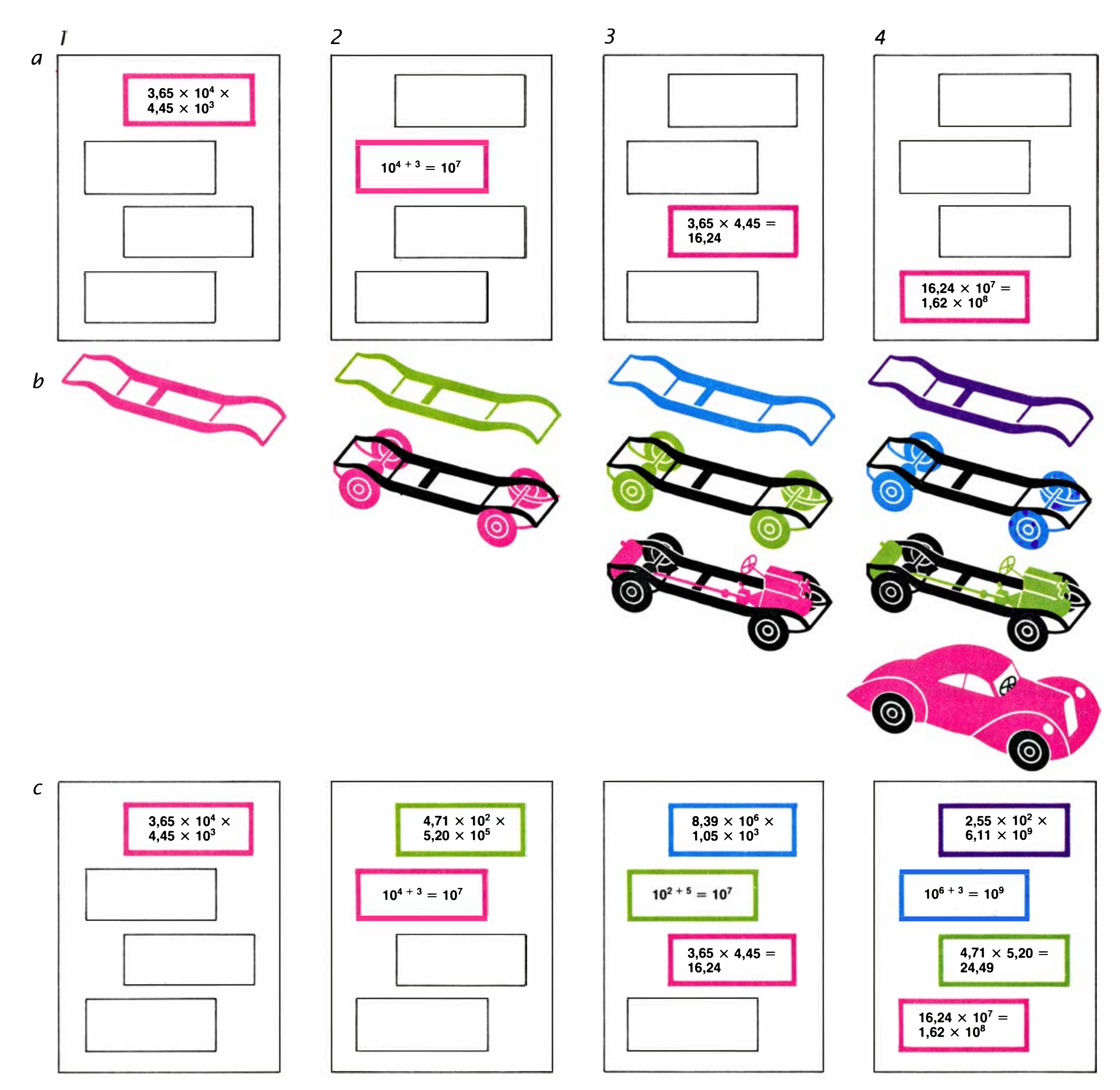
El acceso a la memoria puede también causar "atascos" cuando hay que recuperar de ésta más de un operando antes de ejecutar cierta operación; así ocurre, por ejemplo, en la multiplicación de dos números almacenados en la memoria y el segundo operando no puede recuperarse mientras el canal de acceso a la misma esté ocupado en traer el primero. ¿Cómo diseñar una memoria que posibilite la búsqueda de varios datos al mismo tiempo?

La respuesta, un sistema conocido por "memoria intercalada" ("interleaved"), resulta particularmente idónea para cálculos numéricos que manipulan, de forma sucesiva, grandes conjuntos de números. La memoria intercalada consta de varias unidades de memoria, de ocho o dieciséis, y son accesibles por separado. La primera dirección de memoria (esto es, el lugar donde se aloja el primer número a operar) está en la primera unidad de memoria; la segunda dirección en la segunda unidad, y así sucesivamente. Si hay ocho unidades, la novena dirección volverá a hallarse en la primera unidad; la décima, en la segunda unidad, y así en adelante. En el dominio de este sistema, pueden asignarse, a un mismo tiempo y a través de distintos canales, varias unidades de memoria; asimismo, pueden recuperarse simultáneamente los datos con los que operar, sin tener que esperar canales libres.

1. PROCESAMIENTO PARALELO, en razón del cual una tarea sencilla de computación se divide entre diferentes procesadores (unidades que hacen el trabajo de computación). Lo vemos aquí ejemplificado por este panel de un computador N-CUBO. Cada uno de los 64 cuadrados marrones corresponde a un procesador cuya potencia viene a ser la de un minicomputador VAX 11/750. Los rectángulos negros son unidades de memoria; seis de ellos (que suman medio megabyte de memoria) se dedican a cada procesador. El computador equipado con este panel puede tener de uno a 16 más (es decir, hasta 1024 procesadores), amén de un panel central que encierra, a su vez, un procesador central que distribuye trabajo y datos a los restantes procesadores. Este computador representa un nivel medio en el procesamiento en paralelo. Algunos sistemas paralelos incluyen menos procesadores, aunque más potentes (cierta configuración del Cray X/MP consta de cuatro supercomputadores que se hallan unidos en paralelo), mientras que otros contienen un número mayor de procesadores menos refinados (la "máquina de conexión" comprende 65.536 muy sencillos).

Veamos la segunda fuente de lentitud: el proceso de computación. Como sabe quien haya multiplicado a mano dos números de siete dígitos, las operaciones aritméticas requieren varios pasos elementales. Cuando un procesador multiplica dos números en coma flotante (números que no son necesariamente enteros), debe comenzar por descomponer cada número en un exponente y una mantisa (como en la notación científica); luego, sumará los exponentes, multiplicará las mantisas (operación que suele precisar, ella sola, bastantes pasos) y expresará, por último, el producto en notación científica. La unidad funcional que realiza la multiplicación de números en coma flotante puede segmentarse, de suerte que cada porción ejecute una de las tareas mencionadas. Por el contrario, todos los segmentos menos uno del procesador simple quedan inactivos en cualquier momento que lo consideremos, lo que supone un gran despilfarro de potencia operativa.

Una solución del problema podría presentar rasgos característicos de una cadena de montaje fabril. Una cadena de montaje de coches viene a producir una unidad por minuto. La hazaña no implica reunir operarios empeñados, a la vez, en un mismo coche. Antes bien, existe una división del trabajo en subtareas más sencillas, cumplida cada una de ellas en una estación de montaje distinta. En cuanto el coche parcialmente montado pasa por una estación, el operario asignado a la misma ejecuta la ta-



2. CANALIZACION ("PIPELINING"), técnica para aumentar la velocidad de operación de un solo procesador. Imita el método de trabajo de una cadena de montaje. Determinadas tareas de computación requieren varios pasos sencillos, de cada uno de los cuales se encarga un componente distinto del procesador. En un procesador convencional (*a*), mientras un componente ejecuta su tarea los demás permanecen inactivos. En este ejemplo se trata de multiplicar dos números en coma flotante; los pasos a dar son: extraer de cada número el exponente y la mantisa (en la memoria, los números están codificados en notación científica), sumar los exponentes, multiplicar las mantisas y expresar el producto en notación científica. En una cadena de montaje (*b*), ninguna estación de ensamblaje está inactiva: terminada una fase de la fabricación del coche, y éste pasado a otro eslabón de la cadena, el coche siguiente llega a la primera estación. Se ejecutan así varias operaciones a la vez en coches distintos. Un procesador canalizado (*c*) procede de modo muy parecido: acabada una operación con un par de números, se le presenta otro par para que realice la misma operación con ellos, sin esperar que el primero recorra todas las etapas del cálculo.

rea encomendada: atornillar la tapa de la culata o instalar la radio, por citar dos ejemplos. Es la repetición a gran escala de tareas sencillas lo que confiere a la cadena de montaje su rapidez característica.

Se pueden también construir cadenas de montaje para la aritmética de números en coma flotante. Bautizadas con el nombre de "canalizaciones" ("pipelines"), muestran su eficacia en las aplicaciones que requieren muchas repeticiones de la misma operación aritmética. Supongamos, por ejemplo, que determinado programa exige la multiplicación de muchos pares de números. Admitamos, por mor de sencillez, que cada una de las tareas elementales involucradas en la multiplicación de coma flotante emplea una pulsación del reloj interno del ordenador. En cada pulsación del reloj, se introduce un nuevo par de números en el segmento del procesador que ejecuta la primera tarea. Si la tarea precisa 10 segmentos, entonces al final de la décima pulsación del reloj el producto del primer par de operandos estará listo, casi listo el producto del segundo par, al tercer par le faltarán sólo dos etapas para acabar, y así sucesivamente; el undécimo par habrá entrado en el mismo instante en la canalización y el duodécimo y siguientes se hallarán a la espera.

La cuantía en que la canalización aumenta la velocidad de una operación informática depende de la estructura de dicha canalización. Las hay que se han creado para transferencias y recuperación de datos almacenados en la memoria, para operaciones aritméticas y otras afines. Por ejemplo, si una operación de computación necesita dos o más datos de la memoria, y si su búsqueda tarda unas cuantas pulsaciones del reloj, la estructura canalizadora posibilitaría al procesador iniciar búsquedas de datos en pulsaciones sucesivas, sin tener que esperar la llegada del primer dato desde la memoria para pedir el siguiente.

Cuando una cadena de montaje de automóviles se pone en marcha, transcurre algún tiempo hasta que sale al mercado la primera unidad. Su fabricación ha durado días, aunque el ritmo de producción de dicha cadena de montaje sea de un coche por minuto. Pasemos a la canalización. El intervalo que media entre el comienzo de la primera subtarea del trabajo y la salida del primer producto se denomina "latencia". Este intervalo depende del número de segmentos de la canalización y del tiempo necesario para ejecutar cada tarea. La canalización que posea una gran latencia sólo será eficaz si ha de realizar un número elevado de operaciones idénticas, de igual modo que la cadena de montaje sólo tiene sentido en una fabricación industrial. (Si sólo se produjera una docena de coches cada vez que la cadena se pusiera en marcha, el rendimiento estaría muy por debajo de una unidad por minuto.)

Hay un dominio donde el recurso a la canalización reviste especial interés: el del cálculo vectorial. En el léxico informático, se llama vector a un conjunto ordenado de números independientes. Esos conjuntos reciben el nombre de vectores porque, en geometría, sus respectivos números independientes podrían representar coordenadas: un conjunto de tres números podría representar una posición o una dirección en un espacio tridimensional, mientras que un conjunto de cuatro números podría representar una posición o una dirección en un espacio-tiempo de cuatro dimensiones. En informática, los vectores pueden poseer miles de elementos y no necesitan recibir ninguna interpretación geométrica.

Ocurre con frecuencia que cada elemento de un vector debe someterse a operaciones similares. En una "multiplicación vectorial", por ejemplo, el primer elemento de un vector se multiplica por el primer elemento del otro vector; luego, el segundo elemento de uno por el segundo elemento del otro vector, y así sucesivamente. Tales operaciones son candidatos naturales para el recurso a la "canalización", y el soporte físico diseñado para canalizar operaciones vectoriales constituye un hito importante en la historia de las arquitecturas avanzadas de computación. En varios de estos sistemas, distintas unidades funcionales del procesador procesan a la vez grupos de elementos de vectores. Tales máquinas no son procesadores simples en el sentido genuino del término. Antes bien, representan una etapa intermedia entre las máquinas secuenciales y las de procesadores en paralelo.

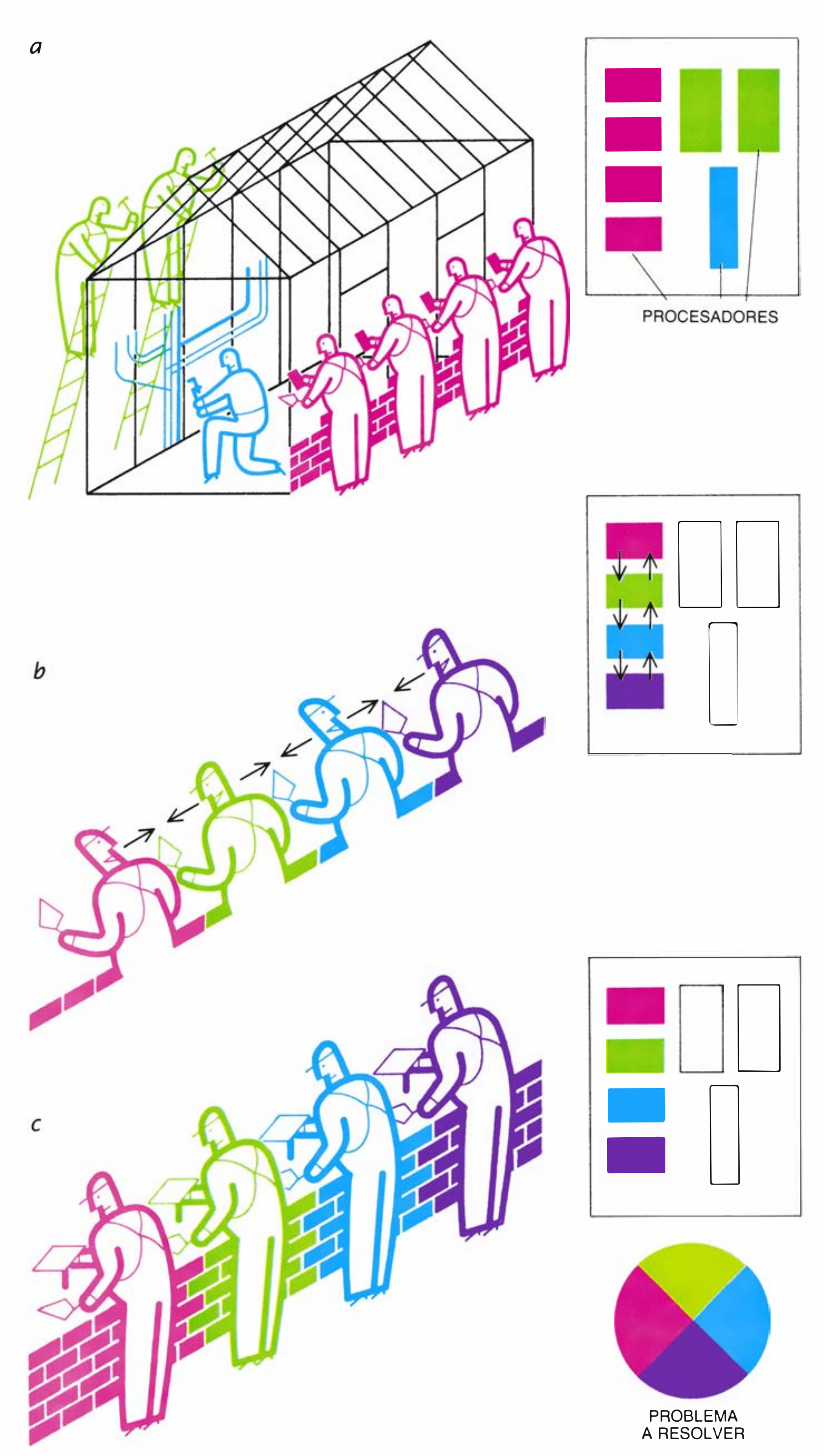
Otra clase de máquinas que podemos asimilar a los computadores paralelos son las máquinas de "instrucción muy larga en una sola palabra" (VLIW, en sus siglas inglesas). En las máquinas VLIW, los procesadores tienen repetidas varias veces cada unidad funcional. Durante cada ciclo, en la etapa en que un procesador convencional toma una instrucción simple de la memoria las máquinas VLIW recogen simultáneamente varias instrucciones. A tal fin, las instrucciones se agrupan juntas en una sola "palabra", que el procesador lee en la memoria. Las instrucciones son transportadas de forma simultánea por las distintas unidades funcionales del procesador; y así, el propio procesador actúa como una máquina que operase en paralelo.

En las máquinas VLIW las operaciones deben jerarquizarse en orden riguroso para evitar determinados conflictos, como el que puede presentarse cuando una operación necesita el resultado de otra. Algunas máquinas VLIW que pueden ejecutar 10 o más instrucciones a la vez acaban de introducirse en el mercado, y se han proyectado ya máquinas capacitadas para acometer cientos de operaciones simultáneamente. La arquitectura de las VLIW resulta especialmente idónea para ejecutar operaciones insólitas, aplicación que las arquitecturas vectoriales y canalizadoras, que exigen cierto grado de regularidad, resuelven con poco rendimiento.

Todos estos sistemas —memoria intercalada, canalización, procesamiento vectorial y maquinaria VLIW— constituyen medios eficaces para aumentar la velocidad y el rendimiento de los procesadores simples. Pero persisten algunas aplicaciones en las que ni siquiera los más rápidos de los procesadores simples resultan suficientemente buenos.

Sea, por ejemplo, un estudio numérico realista de la cromodinámica cuántica (una teoría de la física de partículas elementales); ese trabajo requeriría el cálculo reiterado de unos 100 millones de valores, que vienen a representar las fuerzas de varios campos en la región tetradimensional espacio-temporal en la vecindad de un protón. La velocidad con que las computaciones pueden llevarse a cabo depende, en parte, de la celeridad con que las señales van de una zona a otra del ordenador. La velocidad de la luz en el vacío, la máxima a la cual ciertas señales podrían viajar, es de unos 220 milímetros por nanosegundo (un nanosegundo es una milmillonésima de segundo). Por tanto, aun cuando los componentes de un procesador operasen a una velocidad inimaginable, un procesador simple de reducidas proporciones sólo terminaría algunos miles de millones de instrucciones por segundo, cifra insuficiente para un problema de cromodinámica cuántica. Para tales aplicaciones, la respuesta estriba en enlazar varios procesadores en un auténtico computador en paralelo.

Al proyectar una máquina en paralelo, el ingeniero se enfrenta con un



3. ANALOGIA DE LA CONSTRUCCION DE UNA CASA con determinados aspectos del procesamiento en paralelo. Cada trabajador viene a representar un procesador. En este ejemplo, la cuadrilla entera de operaciones "equivale" al ordenador en paralelo. Igual que los trabajadores de distintos ramos –fontanería, electricidad y albañilería– intervienen de consuno en la construcción de una casa (*arriba, a la izquierda*), así también cabe idear un ordenador que encierre distintos tipos de procesadores, especializado cada uno en una tarea determinada (*arriba, a la derecha*). Otra razón de comparación: los albañiles tienen que ponerse de acuerdo cuando han de levantar los muros para evitar irregularidades (*b, izquierda*); lo propio harán los procesadores de un computador en paralelo (*b, derecha*) cuando han de abordar al alimón un mismo problema enrevesado. La analogía puede extenderse incluso hasta la división del trabajo. En la construcción, cada albañil levanta un segmento de pared (*c, izquierda*); cada operario ejecuta idéntico trabajo, si bien en un lugar distinto. En un problema informático (*c, a la derecha*), se procede también a dividir las tareas de suerte tal que cada procesador acometa la misma operación, aunque con un conjunto distinto de datos (que podrían representar las condiciones físicas de una región particular del espacio). La división es conocida también por descomposición de dominios y se aplica en la teoría física de la cromodinámica cuántica.

amplio abanico de opciones. Por ejemplo, ¿cuántos procesadores debería haber? ¿cómo conectarlos? ¿cuál ha de ser el grado de refinamiento de cada uno? Para mejor entender algunas de las opciones que se le ofrecen al ingeniero de procesadores paralelos, sirvámonos de una analogía. Supongamos de nuevo que el lector ha decidido construirse una casa, pero ahora se lo ha encargado a una empresa. Cada obrero de la constructora representa un procesador individual; la cuadrilla entera se asimila al computador. ¿Qué características definen a este "computador", y en qué difiere de otras formas de organizar múltiples "procesadores"?

El sistema es paralelo: obreros distintos tienen a su cargo diferentes partes del trabajo al mismo tiempo. También es heterogéneo: obreros de diversos ramos (albañiles, electricistas, etcétera) realizan tareas diferentes. Cabe imaginar, no obstante, otras situaciones en las que lo mejor sería emplear varios obreros de un mismo ramo, de igual modo que muchos computadores en paralelo avanzados constan de una batería de procesadores idénticos.

Otro rasgo distintivo de este sistema es el relativo refinamiento de los "procesadores": cada uno posee varias "unidades funcionales" (manos, hombros, piernas, etcétera), capacitadas para operar simultáneamente. Se dice entonces que el sistema tiene un "tamaño grosero", porque a cada obrero se le confía una cantidad sustancial de trabajo; ni resulta fiable su eficacia en una red compleja de comunicaciones entre los procesadores.

Ello no significa que la comunicación entre los obreros sea irrelevante. Los albañiles que construyen un muro han de avanzar en estrecha sintonía para asegurar la uniformidad de la pared, con juntas limpias entre secciones levantadas por distintos obreros. En este caso, la red de conexiones tiene lo que se ha llamado "topología de la frontera contigua": la información se transmite entre operarios contiguos. El sistema en cuestión podría presentar una sobrecarga (esto es, podría comportar que se empleara más tiempo transmitiendo información y menos tiempo desarrollando su labor), si los operarios necesitaran pasar frecuentes mensajes a otros que se hallaran muy distantes. Por ese motivo, hay arquitecturas avanzadas que se apoyan en redes de comunicaciones basadas en topologías distintas del sistema de frontera contigua.

Otra importante consideración atañe

a las instrucciones dadas a los operarios. En el símil de la albañilería, cada obrero desarrolla una tarea ligeramente distinta de la que cumplen sus compañeros (la pared exterior de una casa presenta irregularidades: puertas, ventanas, ornamentos, etcétera); cada operario sigue, pues, un conjunto propio de instrucciones. En una situación en que los trabajadores desempeñen tareas idénticas —por ejemplo en una fábrica textil donde cada uno manejara un telar, y los telares produjeran idénticas muestras— pueden repartirse idénticas instrucciones, y a la vez, a todos los obreros. Ambas situaciones hallan su réplica en las arquitecturas avanzadas de ordenadores.

Un último punto de la analogía con la construcción: la forma en que se han asignado las tareas a los operarios. En lo que concierne a la albañilería, se ha dividido el trabajo en varias subtareas más sencillas, similares unas a otras; cada albañil "resuelve" una porción del problema. Conocido por descomposición en dominios, este modo de dividir resulta apropiado cuando la dificultad principal que plantea un problema reside en el tamaño de la base de datos que se maneja en el problema en cuestión. En nuestro ejemplo, se tarda muy poco en poner un ladrillo, pero mucho en la colocación de todos. Cada obrero se responsabiliza, pues, de un determinado montón.

Tales problemas son muy comunes en computación —el modelizado de la cromodinámica cuántica es un buen ejemplo— y la descomposición en dominios es la forma más natural de resolverlos. Es importante en la descomposición en dominios asegurarse de que la red mediante la cual son conectados dé muestras de algún parecido con la topología natural del problema que se está resolviendo. Aquí la red de frontera contigua se acomoda bien con la geometría real de una casa.

La solución de algunas tareas no pasa por la descomposición en dominios. Se abordan mejor mediante el método de descomposición funcional. La compañía constructora constituye, de nuevo, un buen punto de comparación. En la empresa hay directivos, administrativos y obreros; entre éstos, fontaneros, electricistas y albañiles. El trabajo de la compañía se reparte en subtareas adecuadas a la cualificación profesional de cada grupo. Un computador en paralelo que se base en la descomposición funcional puede a veces trabajar bastante más deprisa que un computador secuencial, aunque el grado de aceleración del proceso informático depende del número de tareas distintas que cubre el problema a resolver. Lo habitual es la combinación de ambos métodos: la descomposición funcional y la descomposición en dominios; tornando al ejemplo, la obra se reparte entre albañiles, fontaneros y electricistas y la subtarea de albañilería se fracciona entre varios obreros.

La metáfora de la empresa constructora aporta una buena exposición de los puntos implicados en la creación de un computador en paralelo. Detengámonos en algunas de las cuestiones abordadas por los ingenieros y la forma en que han decidido resolverlas.

Ante la oportunidad de incluir en el sistema más de un procesador, nada nos obliga a proyectar un sistema homogéneo escogiendo los idénticos. ¿Por qué no combinar una mezcla apropiada de procesadores especializados en distintos cometidos y crear un sistema heterogéneo? Después de todo, muchas organizaciones humanas siguen ese camino. Por citar un ejemplo, la compañía aérea que cubre, además de ciudades próximas, territorios extensos, desplegará, de acuerdo con cada necesidad, aviones de hélice, aviones de tamaño medio y grandes reactores.

Se han fabricado ya algunos computadores paralelos; y hay muchos más proyectados. Las máquinas construidas suelen contar con dos clases de procesadores: los de empleo general y los específicos de la aritmética de coma flotante. Se han diseñado otros sistemas que combinarían procesadores especializados con arquitecturas idóneas para inteligencia artificial, gráficos, bases de datos y procesamiento numérico, entre otros objetivos; si bien, los grandes sistemas no han alcanzado todavía la fase comercial. Habida cuenta de que los computadores en paralelo homogéneos predominan entre las máquinas puestas en servicio, centraremos el resto del artículo en los sistemas homogéneos.

Hemos elegido ya la batería de procesadores. ¿Cómo le indicamos a cada procesador qué pasos debe seguir durante la computación? En la empresa constructora, cada operario cumplía un repertorio cerrado de instrucciones. En terminología informática, ese sistema se conoce por corriente de instrucciones múltiples/corriente de datos múltiples (o MIMD, de sus siglas en inglés). Mediante un sistema MIMD se puede descomponer un programa en segmentos, distribuir las tareas y los datos asociados entre los procesadores y dejar que éstos trabajen con bastante independencia. Hay otra opción: la corriente de instrucción única/corriente de datos múltiple (o SIMD, de sus siglas en inglés). En un sistema SIMD, se remite a todos los procesadores el mismo repertorio de instrucciones. En cada paso de la computación, cada procesador ejecuta, a la vez, la misma orden que todos los demás, o permanece inactivo.

Una buena analogía con los sistemas SIMD es el juego del bingo, donde se van cantando los números uno detrás de otro. Cada número constituye una instrucción: "Busque en su cartón el número que he cantado. Si el número aparece en el cartón, márquelo. Si no aparece, no haga nada." Esta orden determina que un jugador se quede quieto o que realice la misma acción que otros participantes. Los números de su cartón son los datos con que cuenta el jugador. En cierto modo, el cartón puede asimilarse a la unidad local de memoria: una unidad de memoria disponible para un solo procesador. Si el locutor tiene una voz recia, podrían jugar, casi al unísono, bastantes personas. La acción entera se desarrolla en las unidades de memoria, o en su vecindad, cada una de las cuales está asociada a un procesador de instrucciones distinto (cada jugador).

Una aplicación natural de las máquinas SIMD es el problema de la cromodinámica cuántica, antes mencionado. Para resolverlo, cada procesador podría calcular la intensidad de varios campos en una región particular del espacio. Cada procesador tendría entonces que ejecutar esencialmente la misma operación con su propio conjunto de datos. Entre las aplicaciones habituales de una estructura SIMD citaremos, además, el procesamiento de imágenes, el análisis de flujos y las búsquedas a través de grandes bases de datos.

Se han construido diversos sistemas dotados con la estructura SIMD, y han demostrado su rendimiento en tareas numéricas y simbólicas. A modo de ejemplo suele aducirse el "Goodyear Massively Parallel Processor", que alberga 16.384 procesadores, y la "Connection Machine", con 65.536 procesadores, de extremada sencillez cada uno. A pesar de lo cual, las máquinas realizan computaciones generales, además de las tareas propias de una estructura SIMD.

¿Cómo conectar los procesadores de un computador paralelo, entre sí y con las unidades de memoria? El sistema de interconexiones en virtud del cual los procesadores comparten la información entre sí y con las unidades de memoria constituye una de las peculiaridades más importantes de cual-

quier sistema paralelo; se trata, también, del dominio donde aparece la mayor diversidad de los sistemas paralelos. En el método de "memoria compartida", todos los procesadores están conectados a un sistema común de memoria, ya sea por conexión directa, por una red de conexiones o por un "bus de memoria". El "bus" viene a ser un canal por cuyo cauce los distintos procesadores pasan demandas de información y por donde la unidad de memoria remite los datos a aquéllos.

En un sistema de memoria compartida, cada registro de datos se halla al alcance directo de todos los procesadores; si uno necesita determinados datos para su trabajo, los lee de la memoria. Puesto que otros procesadores pueden estar modificando continuamente los datos almacenados, hay que tener cuidado de que ningún procesador lea en un registro de la memoria antes de que se haya depositado allí el valor apropiado. La coordinación necesaria corre a cargo del soporte lógico, pero no está de más alguna ayuda del soporte físico. Seré más explícito: en algunos sistemas, quedan reservadas determinadas zonas de la memoria para elementos que pueden guardar pistas de la inscripción de ciertos registros, y así proteger de intentos prematuros de acceso a los datos que dichos registros almacenarán.

La memoria compartida presenta una limitación: podría resultar difícil o caro crear una memoria que subviniera a múltiples procesadores a la vez. Si cada procesador posee una conexión directa con la memoria, hay 1000 procesadores y la memoria se descompone en 1000 bancos distintos para evitar los "embotellamientos", se precisará un millón de conexiones. Por otro lado, si todos los procesadores tienen acceso a la memoria a través de un "bus" común, entonces éste debe alcanzar suficiente velocidad para satisfacer las demandas de memoria que urjan todos los procesadores simultáneamente. A este respecto, el proyecto de una memoria común recuerda la regulación del tráfico en una autopista de accesos limitados. Para cierto límite de velocidad y determinado número de carriles, cada autopista tiene una capacidad de coches por minuto. Cuando intentan usar la autopista más coches de los que la capacidad de ésta permite, el tránsito se hace más lento y queda congestionado. Aunque una solución sería construir autopistas con más carriles (o varias autopistas paralelas), lo cierto es que existen límites de orden práctico para ello.

Cabría otra posibilidad: construir una red de comunicaciones similar a la red telefónica mundial. Hay cientos de millones de receptores en el mundo y, salvo contadas excepciones, cada teléfono puede conectarse con cualquier otro. No se trata, evidentemente, de ir tendiendo hilos entre cada par de receptores. Antes bien, grupos de aparatos están conectados a distintos sistemas de conmutación. Los sistemas de conmutación se conectan entre sí de una manera jerárquica, de suerte que cualquier estación conmutadora es responsable de un número manejable de conexiones. Cuando se solicita una conexión (marcando un número de teléfono), la llamada se propaga a través de distintas estaciones de conmutación hasta llegar al receptor deseado.

Pero las conexiones reticulares aumentan la latencia de acceso a memoria, ya que cada demanda de información debe atravesar diversas fases de la red en su camino hacia la unidad de memoria. Además, igual que sucede en

4. EN LAS ARQUITECTURAS DE MEMORIA COMPARTIDA, los procesadores de un sistema paralelo tienen acceso a unidades de memoria comunes. En el modelo elemental de memoria compartida (*a*), cada procesador se halla conectado directamente a cada banco de memoria. Este sistema presenta el inconveniente de que cada procesador o banco de memoria debe asistir a un gran número de conexiones. Una opción alternativa (*b*) la encontramos en el "bus": una línea de comunicación común a lo largo de la cual los procesadores pasan demandas a los bancos de memoria y éstos devuelven los datos. El bus puede quedar sobrecargado cuando han de transitar un número elevado de mensajes. Cabe recurrir también a la red omega (*c*); en ésta, los procesadores se conectan a las memorias mediante cajetines commutadores, cada uno de los cuales ofrece dos líneas de llegada y otras dos de salida. En una red omega cada procesador tiene una línea directa a cada memoria, pero no se precisan tantas líneas de comunicación como en el esquema de conexiones directas. La ventaja se torna más palpable a medida que crece el número de procesadores y bancos de memoria. El inconveniente es que el mensaje debe atravesar múltiples estaciones conmutadoras.

la red telefónica, puede haber períodos de actividad punta, cuando la demanda de conexiones satura la red, dando como resultado demoras muy largas. Se han elaborado varios tipos de red para relajar la congestión manteniendo un coste razonable y la velocidad de acceso a la memoria.

Si tan difícil y caro es construir máquinas de memoria compartida, ¿por qué no asignar cierta cantidad de memoria a cada procesador y conectar los pares procesador-memoria (o nodos, como se les denomina en este contexto)? Los sistemas que se han ideado según esa línea tienen memoria local (conocida también por memoria distribuida). En arquitecturas de memoria local, los datos deben repartirse por los nodos apropiados al comienzo de la computación. En el curso de la misma será necesario que los nodos recaben datos almacenados en las memorias de otros nodos. Lo que se consigue enviando un mensaje al nodo correspondiente en el que se solicite que la información le sea entregada por medio de un mensaje de respuesta. Si los dos nodos implicados no están conectados directamente, otros nodos intermedios pueden vehicular el mensaje.

Se han pergeñado varios esquemas para la conexión entre nodos de los sistemas de memoria distribuida. El ingeniero de esos sistemas se propone varios objetivos: rápida comunicación entre nodos, ningún nodo debe soportar demasiadas conexiones y la topología de las conexiones ha de ajustarse a la geometría natural del problema a resolver. El análisis de la cromodinámica cuántica proporciona un ejemplo adecuado para esta última condición. Si cada nodo calcula los valores de varios campos de un determinado volumen del espacio, entonces debería conectarse a los nodos responsables del cálculo de esos mismos valores en los volúmenes circundantes.

Uno de los esquemas de interconexión más sencillos lo encontramos en el anillo. En éste, cada nodo se conecta a otros dos y la línea de conexiones dibuja un círculo. Otro esquema bastante elemental es la malla, donde cada nodo está conectado a sus cuatro nodos vecinos más próximos. Aunque varios nodos se conecten en anillo o en malla, el número de conexiones que cada procesador debe sostener es pequeño (dos o cuatro, respectivamente). El procesador asociado a cada nodo accede a su propia memoria a una velocidad apropiada, y los canales de comunicación cursan sólo entre unos cuantos nodos.

Estas topologías simples adolecen de algunos inconvenientes. El número de nodos involucrados en el envío de un mensaje de un nodo a otro puede ser muy grande (en un anillo, hasta la mitad de los nodos podrían hallarse implicados), dando como resultado largas demoras. Si un nodo no tiene con qué entretenerse mientras espera la llegada de datos, su capacidad estará desaprovechada y si hay que recurrir a la demanda frecuente de datos de otros nodos, se malgastará una cantidad notable de la potencia del ordenador.

Para paliar este problema se ha pensado en una topología ligeramente más elaborada: el esquema n-cúbico o hipercubo. Este método conecta los nodos como si residieran en los vértices de un cubo multidimensional. Así, habría ocho nodos conectados en un cubo de tres dimensiones y 16 nodos en la disposición topológica que se crearía con un cubo tetradimensional. (Los cubos de una y dos dimensiones son, respectivamente, la línea recta y el cuadrado.) En general, un cubo p-dimensional puede conectar 2^p nodos. El camino más largo que necesitará recorrer un mensaje es el logaritmo en base 2 de p. Hoy día existen en uso sistemas comerciales que enlazan 2^{12} nodos mediante una topología hipercúbica.

Ante las diferencias que hay entre las topologías de los sistemas de memoria compartida y los de memoria local, y consideradas también las distintas potencias de las dos, nada más natural que intentar combinarlas. En ese sentido, podríamos imaginar un sistema híbrido en el que cada nodo constara de algunos procesadores que compartieran una memoria. Los nodos, o "racimos" en este caso, podrían conectarse mediante topologías de malla o de hipercubo. El inverso de este sistema, en el que nodos que poseen cada uno memoria local comparten también una memoria común, se encuentra ya en algunos sistemas.

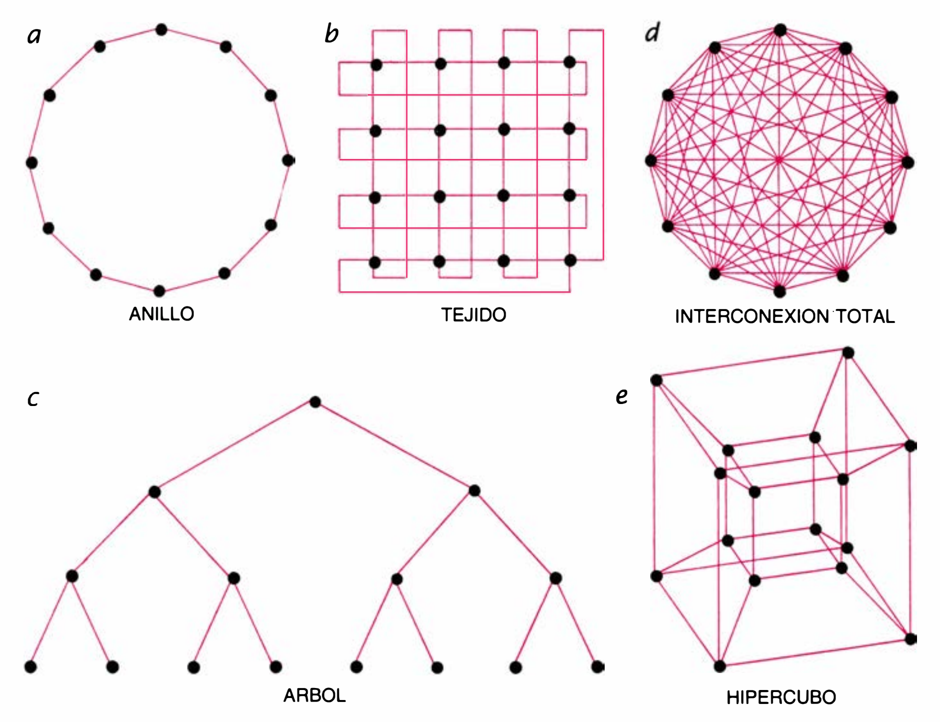
Así como se está experimentando con el número de procesadores de un ordenador y la morfología de la conexión entre dichos procesadores, ¿por qué no experimentar también con la forma con que se controla la operación de la máquina? Casi todos los computadores construidos hasta ahora, incluidos los computadores en paralelo, están controlados por programas. El programador redacta un conjunto de instrucciones, que se traducen luego a lenguaje de máquina: el código que describe los pasos reales de la computación que debe recorrer la máquina. El programa establece, con precisión y sin ambages, el flujo de trabajo a través del computador.

Las arquitecturas en paralelo permiten imaginar formas radicalmente diferentes de controlar el trabajo del computador. Pero en su mayoría se hallan todavía en fase experimental. Empecemos por el método de "flujo de datos". En una arquitectura de ese tenor, se representa el problema mediante un grafo dirigido: un conjunto de puntos, llamados nodos, unidos por arcos. Los nodos simbolizan las operaciones matemáticas o lógicas y los arcos representan el flujo de datos de una operación a otra. En la máquina, los nodos podrían representarse por pares de memoria-procesador; los arcos, por conexiones físicas entre los procesadores.

En una máquina de flujo de datos, el nodo no ejecuta sus operaciones en una secuencia predeterminada y especificada por un programa. Espera, por contra, hasta que ha recibido todos los datos requeridos para cumplir su misión; sólo entonces ejecuta la operación y pasa la respuesta al nodo siguiente. Eso mismo ocurre en una carrera de relevos, en la que el segundo corredor arranca en cuanto recibe el testigo de su compañero. En el caso de una máquina de flujo de datos, no obstante, un procesador puede esperar resultados de varios procesadores más, y no necesariamente de uno solo.

Podemos imaginar un sistema de flujo de datos en el que un número elevadísimo de nodos ejecute instrucciones que no guarden relación alguna con las que la mayoría de los otros nodos están cumpliendo. Puesto que cada uno aborda la tarea asignada cuando dispone ya de los datos necesarios, no hay peligro de que se precipite y recupere datos de un registro de memoria antes de que el valor de última hora se haya almacenado allí. Tampoco se corre peligro de interferencia con otros procesadores, en el sentido de pretender leer en el mismo registro de la memoria a un mismo tiempo. Ni se produciría la situación conflictiva de que dos mensajes se aprestaran a utilizar el mismo canal de comunicación simultáneamente. En resumen, el sistema de flujo de datos evita, pues, la sobrecarga más frecuente en las arquitecturas en paralelo.

A esta solución contrapónese la arquitectura de demanda dirigida, también posible. Las máquinas de demanda dirigida se parecen a las de flujo de datos en su representación plástica: un conjunto de nodos donde se ejecutan las operaciones y arcos para el curso de los datos. Se distinguen en que se realiza la operación en un nodo sólo cuando otro nodo necesita el resultado



5. ARQUITECTURAS DE MEMORIA DISTRIBUIDA. Establecen conexiones entre procesadores de suerte que cada uno posee absoluto control sobre parte de la memoria. Dos de los modelos más sencillos son el anillo (*a*) y la malla (*b*). La arquitectura del árbol binario (*c*) reviste mayor especialización, y sirve, sobre todo, para los sistemas expertos, que se apoyan frecuentemente en procesos de toma de decisiones de morfología dendriforme. Aunque podríamos conectar cada procesador con cada uno de los demás (*d*), semejante diseño requiere un excesivo número de conexiones. Una arquitectura innovadora que se está introduciendo con decisión es la topología hipercúbica (*e*); en ella, los procesadores se conectan como si se hallaran en los vértices de un cubo multidimensional. En este sistema, cada procesador envía mensajes a cualquier otro de los restantes por una ruta bastante corta, y los procesadores no necesitan sostener excesivas conexiones.

de la primera. Las expresiones se evalúan sólo si se exige su resultado; en las arquitecturas de flujo de datos, por contra, se ejecuta cada operación tan pronto como es posible, al margen de que se necesite o no su resultado.

Las arquitecturas de demanda dirigida implican, por su propia naturaleza, una demora de tiempo: el que invierte un mensaje en pasar de un nodo a otro "superior" para comunicarle a este último que ejecute su operación. Las arquitecturas de flujo de datos y las de demanda dirigida tienen en su favor la posibilidad de una programación más eficiente que con las máquinas en paralelo convencionales. Bastaría con que el usuario especificara las operaciones a realizar por cada nodo y el flujo de datos entre nodos, sin preocuparse por la mecánica exacta o la temporización del algoritmo a seguir.

Los científicos interesados en la informática tienen una experiencia importante, aunque limitada, de las arquitecturas avanzadas. Algunas arquitecturas paralelas con un futuro halagüeño (pensemos en las máquinas SIMD, la "Connection Machine", el Hipercubo del Instituto de Tecnología de California, así como varias máquinas de memoria compartida) han sido objeto de estudio y se las ha reputado viables en una amplia gama de aplicaciones. No se pierda de vista, además, que los soportes físicos y lógicos hoy disponibles se hallan en una fase rudimentaria. Son pocas las comparaciones cuantitativas a fondo de máquinas diferentes. Por tanto, aunque prevemos el enorme interés que las máquinas en paralelo despertarán el día de mañana, con unas prestaciones muy por encima de las ofrecidas por las máquinas secuenciales, no nos vemos capacitados para aventurar qué arquitecturas específicas o qué niveles de rendimiento se alcanzarán en un futuro más o menos inmediato.

El procesamiento en paralelo ha avanzado tanto merced, en buena parte, a la visión de futuro mostrada por el Gobierno y la industria que ha subvencionado muchos proyectos universitarios, cuyo resultado es ese impresionante número de esfuerzos comerciales que cubren múltiples arquitecturas. Vemos un prometedor futuro a corto plazo para los nuevos soportes físicos. No aventuraríamos lo mismo en lo concerniente a los soportes lógicos: no se ve esa decidida inversión orientada hacia el sector, crucial aunque menos espectacular. Se ha tardado más de 30 años en crear la base actual de soportes lógicos secuenciales. Los avances en las técnicas de soporte lógico podrían lograr la aplicación de esta base lógica en los últimos prototipos de máquinas; pero si se trata de dar saltos cualitativos habrá que empeñar mucho más y replantear y reescribir las aplicaciones fundamentales en la industria, los servicios y la investigación.

Un paso crucial será el establecimiento de algunas normas. Empleados en máquinas secuenciales, lenguajes como FORTRAN, COBOL y LISP aportaron los cimientos para ciertos trabajos científicos fundamentales y otros relacionados con el mundo de los negocios y la inteligencia artificial. Nadie niega que eran lenguajes imperfectos, pero fueron lo bastante buenos para permitir el desarrollo de los fundamentos de soporte lógico secuencial que pudieron desplegarse a medida que crecía el refinamiento del soporte físico. El mundo de las arquitecturas en paralelo carece todavía de un entendimiento y acuerdo similares en lo relativo a normas equiparables de soporte lógico.

El futuro vendrá conformado por la nueva tecnología. Los progresos en el dominio de los semiconductores, y quizas en la técnica basada en los nuevos superconductores de altas temperaturas, promoverán nuevas inquietudes en el diseño del soporte físico. Habrá seguramente nuevas máquinas híbridas cuyas interconexiones ópticas enlazarán procesadores digitales de alta velocidad. Avances que incluirán los modelos esenciales de computación que hemos descrito aquí.

Más radical es la posibilidad de computación enteramente óptica, que permitiría computaciones analógicas rápidas de distintos algoritmos que hoy sólo se realizan digitalmente de manera grosera y consumidora de tiempo. Ignoramos si tales computadores ópticos se limitarán necesariamente a algunas áreas especializadas de procesamiento de señales o si podrán desarrollarse en máquinas de aplicación general.

Es obligado mencionar, por último, el trabajo reciente en el campo de los modelos de redes neuronales. Su atractivo es obvio: proporcionan un modelo informático que quiere compararse al del cerebro humano. Las máquinas de red neuronal son potencialmente muy superiores a los computadores convencionales en el reconocimiento de patrones, aunque inferiores en la aritmética de coma flotante y otras tareas monótonas. Los investigadores han construido ya algunos soportes físicos de redes neuronales sencillas. Es de esperar que este nuevo modelo de computación conozca una rápida evolución en el soporte físico y en los algoritmos.

Microcircuitos para computación avanzada

En 1959, en una pastilla de silicio se alojaba un solo transistor; hoy se sobrepasa el millón. En el año 2000 habrá microcircuitos con mil millones de componentes

James D. Meindl

No es aventurado afirmar que el acontecimiento más importante de los últimos 50 años sea el advenimiento de una sociedad cuyo producto principal es la información. La fuerza motriz de esa transformación ha sido el desarrollo de la tecnología electrónica, en especial del circuito integrado fabricado sobre una diminuta oblea de silicio. Desde la llegada del circuito integrado, en 1959, el número de transistores que pueden comprimirse sobre la superficie de esas pastillas ha pasado de la unidad a varios millones. Resultado de ello ha sido que las prestaciones de los circuitos integrados se han multiplicado por un factor superior a 10.000. Y no sólo eso: mientras los microcircuitos experimentaban sensacionales perfeccionamientos y mejoras en su rendimiento y complejidad, su costo permanecía sensiblemente invariable.

Al reflexionar sobre tan formidable progreso tecnológico, la pregunta que brota espontáneamente es la de cuánto tiempo va a continuar. Podemos acotar la cuestión inquiriendo cuándo se dispondrá en el mercado de un microcircuito perteneciente a las fases iniciales de la GSI (por *gigascale integration*, integración a escala giga, esto es, del millar de millones). Para determinar tal fecha es preciso analizar los factores que limitan la integración a escala tan enorme. Tales factores se distribuyen en una jerarquía de cinco niveles de limitación; a saber: naturaleza fundamental, material, dispositivos, circuitería y sistema. Cada nivel de la jerarquía puede abordarse desde tres puntos de vista: el teórico, el práctico y el de la analogía histórica.

Las limitaciones de naturaleza teórica se fundan en principios científicos sólidamente establecidos. Constituyen, en efecto, límites intrínsecos del diseño de dispositivos GSI. Los límites prácticos, por el contrario, no dependen de principios científicos, sino de los procesos de fabricación y del equipo disponible. Las limitaciones de carácter práctico contribuyen a dictar lo que es posible lograr en un momento dado en el marco de una tecnología comercialmente disponible. La analogía histórica puede contribuir a iluminar el futuro a largo plazo de la revolución informática, por comparación con la Revolución Industrial en su fase avanzada. Tomados en conjunción, esos enfoques nos ayudarán a responder la pregunta de cuándo se alcanzará la GSI. Una vez obtenida la respuesta, otearemos, más allá de los límites hoy estimados, los dispositivos en proceso de desarrollo en los laboratorios y que tal vez puedan desencadenar la fase siguiente de la revolución informática.

En este artículo abordaremos primero la perspectiva teórica, nos ocuparemos luego del enfoque práctico y, por último, de la contribución que por analogía pueda aportarnos la historia. Desde la perspectiva teórica, las limitaciones fundamentales emanan de las leyes de la naturaleza, inalterables. Tales restricciones son independientes, por definición, del material que se utilice para construir un dispositivo, del diseño del propio dispositivo, del circuito en el que vaya a utilizarse y del sistema del que forme parte el circuito. Las limitaciones fundamentales constituyen el primero (y el más general) de los niveles de la jerarquía anteriormente mencionada. Como tales, constituyen límites extremos, más allá de los cuales yacen sólo zonas prohibidas al funcionamiento.

Las limitaciones fundamentales proceden de diversos campos de las ciencias físicas. Así, la termodinámica enseña que los niveles energéticos de los átomos y electrones del semiconductor cristalino con que se fabrican los transistores experimentan fluctuaciones estadísticas aleatorias. Ahora bien, los niveles de energía de los electrones se hallan en estrecha relación con la señal que está procesando el transistor. Si tal señal no es considerablemente más intensa que las fluctuaciones aleatorias, éstas podrían confundirse con la señal. Las fluctuaciones imponen, por consiguiente, un mínimo a la energía de conmutación. Dado que las fluctuaciones aumentan con la temperatura, la energía mínima requerida para la conmutación del dispositivo crece con la temperatura. Otras limitaciones fundamentales vienen determinadas por la mecánica cuántica, la teoría electromagnética y la velocidad de la luz.

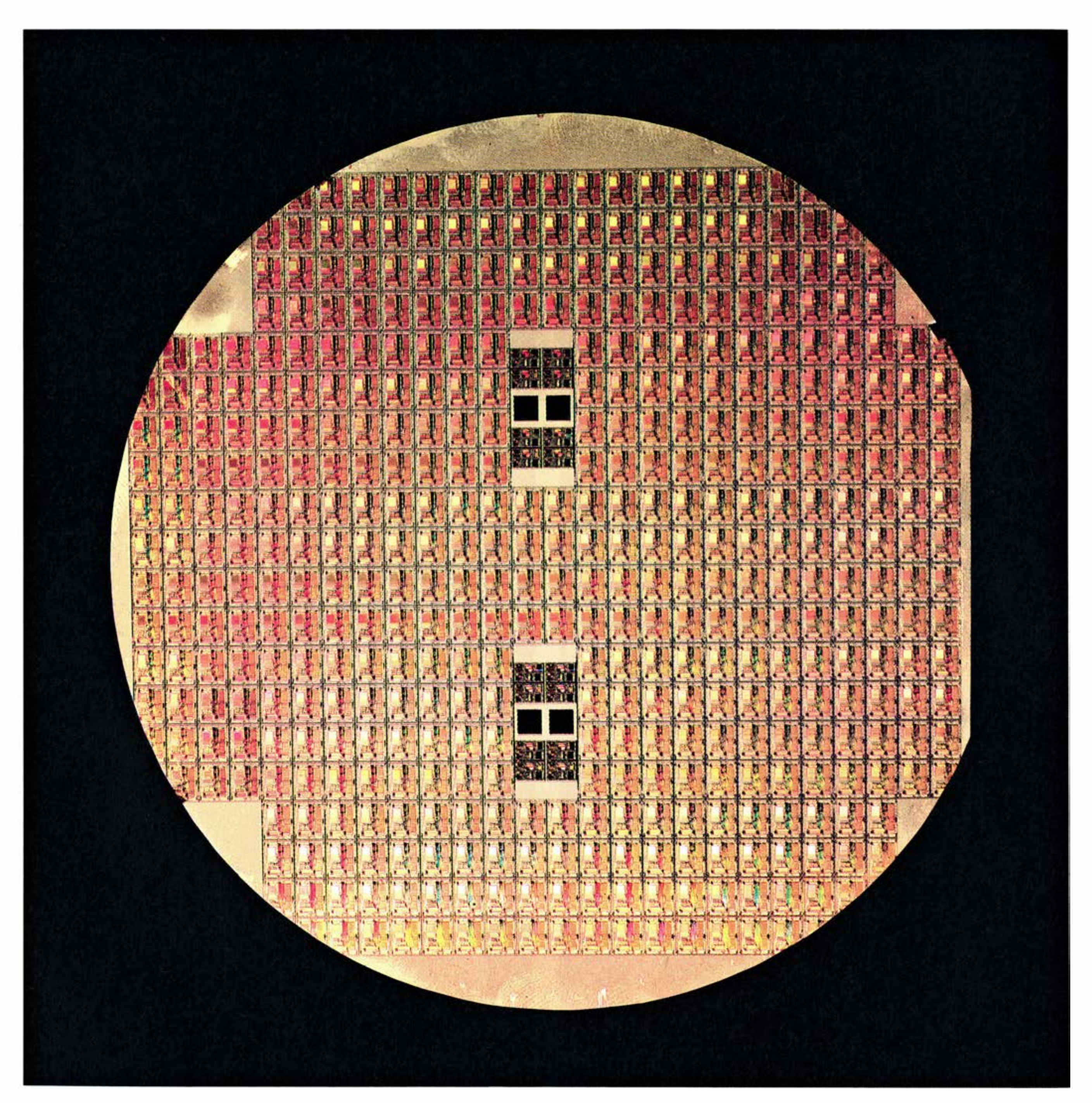
Las limitaciones debidas al material dependen de la estructura y composición química de la sustancia, pero no de la configuración del dispositivo propiamente dicho. En el caso de los circuitos microelectrónicos, los materiales de máximo interés son los semiconductores. Los átomos que componen el retículo cristalino de un semiconductor están ligados entre sí mediante electrones de valencia; cada átomo de silicio tiene cuatro electrones de esos. En condiciones ordinarias, los electrones de valencia están fuertemente ligados a los átomos que componen la red cristalina.

Tal situación puede alterarse por medio del "dopado", esto es, por adición de átomos de impurezas, cuyo número de electrones de valencia es mayor, o menor, que el silicio. El arsénico, uno de los agentes contaminantes más comunes, porta un electrón de valencia más que el silicio. Tal electrón supernumerario suele moverse con libertad por el retículo cristalino. Al aplicar un campo eléctrico al material, los electro-

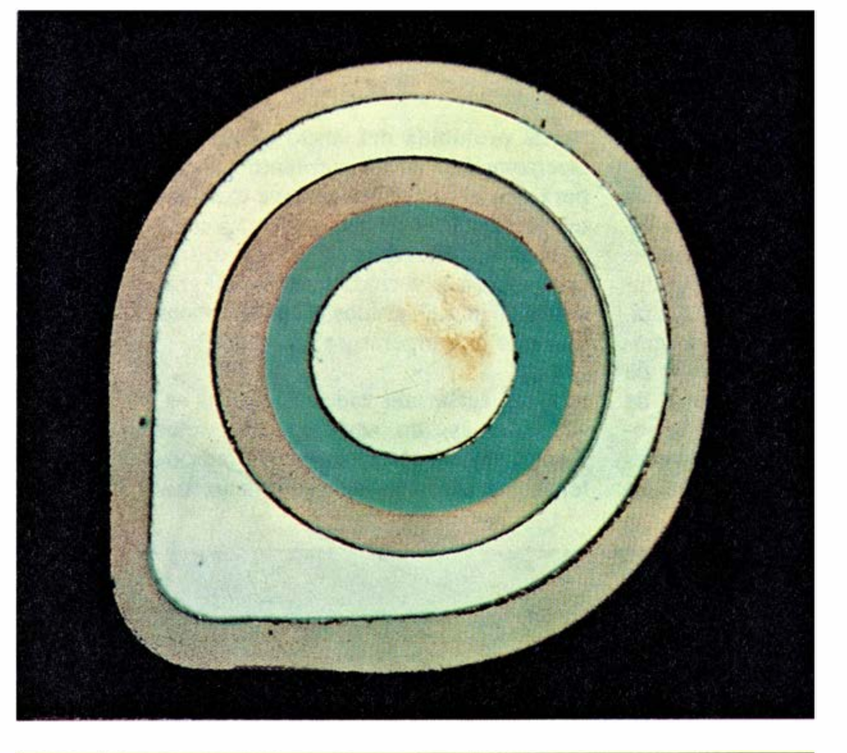
nes libres, portadores como son de carga negativa, conducen una corriente eléctrica. Los dopantes como el boro, por el contrario, tienen un electrón de valencia menos que el silicio. El "hueco" resultante de ese déficit actúa como portador de corriente cargado positivamente. La adecuada elección del tipo y concentración del contaminante posibilita el control preciso de las propiedades eléctricas que exhibirá el semiconductor.

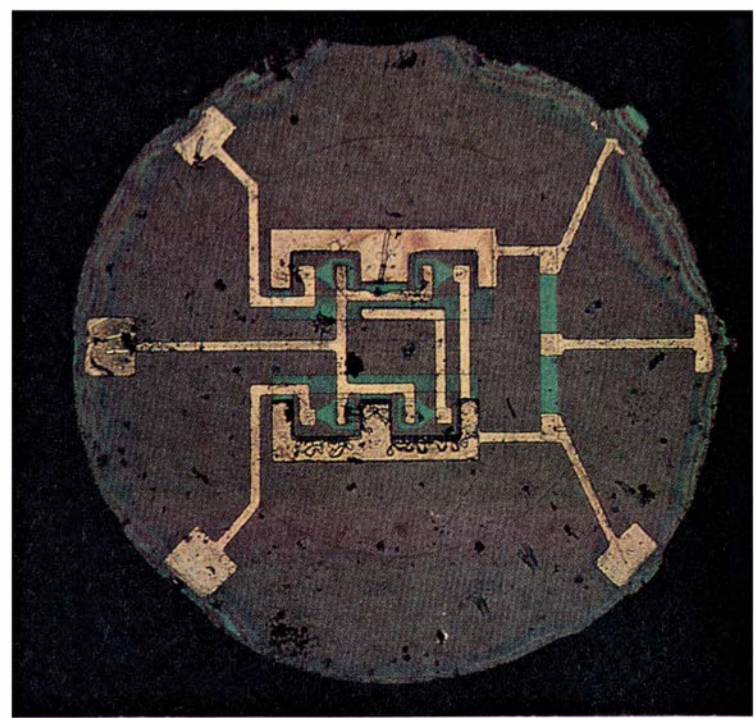
Son varias las propiedades del silicio que lo hacen excepcionalmente idóneo para su empleo en circuitos integrados. Citemos, en primer lugar, la amplitud de su banda prohibida, esto es, la diferencia de energía entre los electrones de valencia y los de conducción. Cuando la banda es demasiado estrecha, basta una ligera elevación de la temperatura para bombear a la banda de conducción un número excesivo de electrones, lo que dificultará o impedirá el control preciso de las propiedades eléctricas del dispositivo. La banda prohibida del silicio es de 1,12 electronvolt, suficientemente ancha para que el material conserve excelentes propiedades de semiconducción en un amplio intervalo de temperaturas centrado en la vecindad de los 300 grados Kelvin (27 grados Celsius, pocos más que la temperatura ambiental).

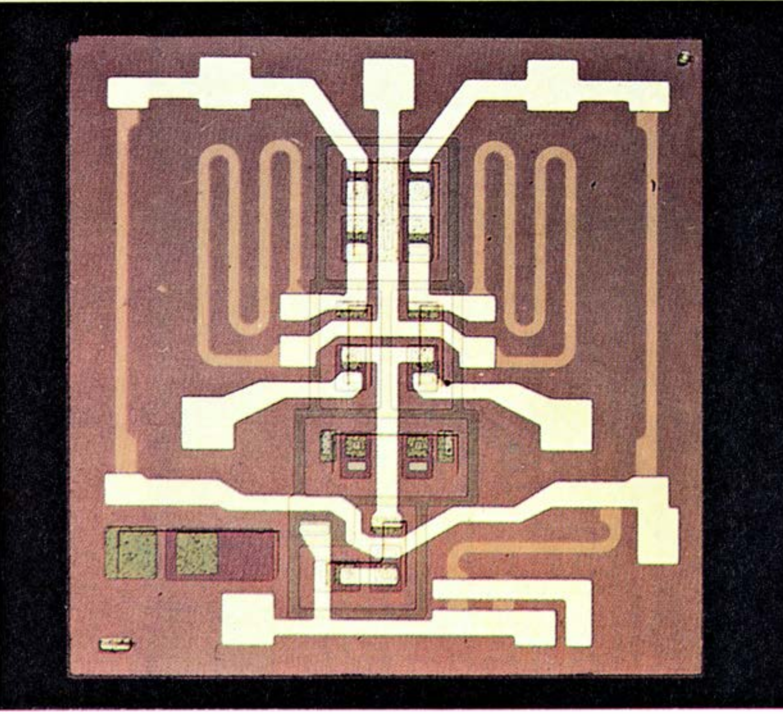
Y no acaba ahí todo. El silicio es, además, un semiconductor elemental abundante, del que pueden obtenerse cristales casi perfectos a un cos-

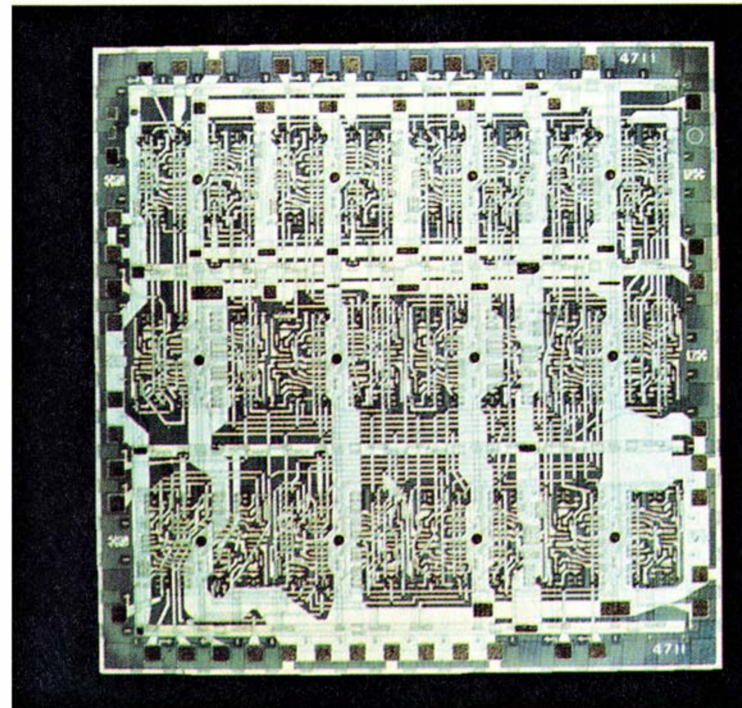


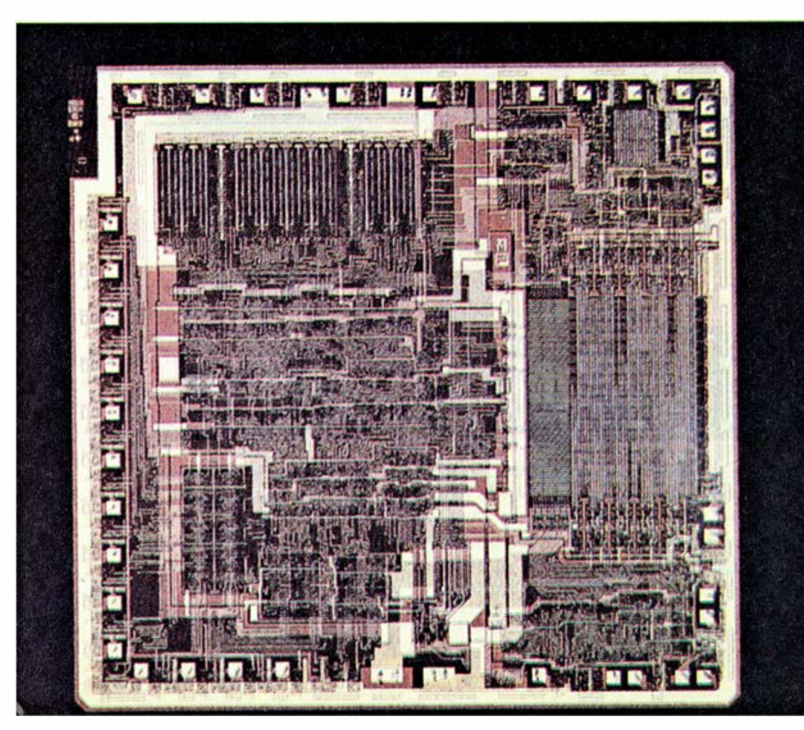
1. OBLEA DE SILICIO que contiene 470 microcircuitos de ordenador, en las fases finales de su fabricación. La oblea, mostrada aquí a escala natural, es una fina rodaja de una "barra" de silicio. Los microcircuitos se fabrican todos a la vez sobre la oblea; los rectángulos próximos al centro son circuitos de prueba, que se utilizan para supervisar los procesos de fabricación. Tras ensayarlos, para detectar los defectuosos, los microcircuitos se recortan con una sierra de diamante, procediéndose al encapsulado de los que funcionan correctamente. Los microcircuitos aquí mostrados son del tipo PACE1750A, que fabrica la Performance Semiconductor Corporation. Cada uno constituye la unidad central de un miniordenador. Se utilizan en aviónica militar y en otros sistemas.

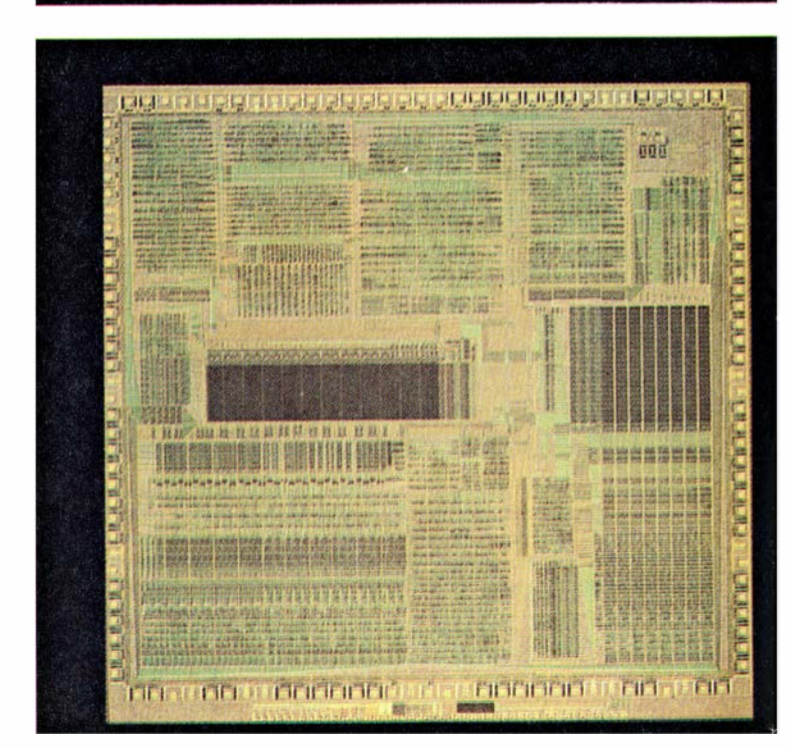







4711
4711

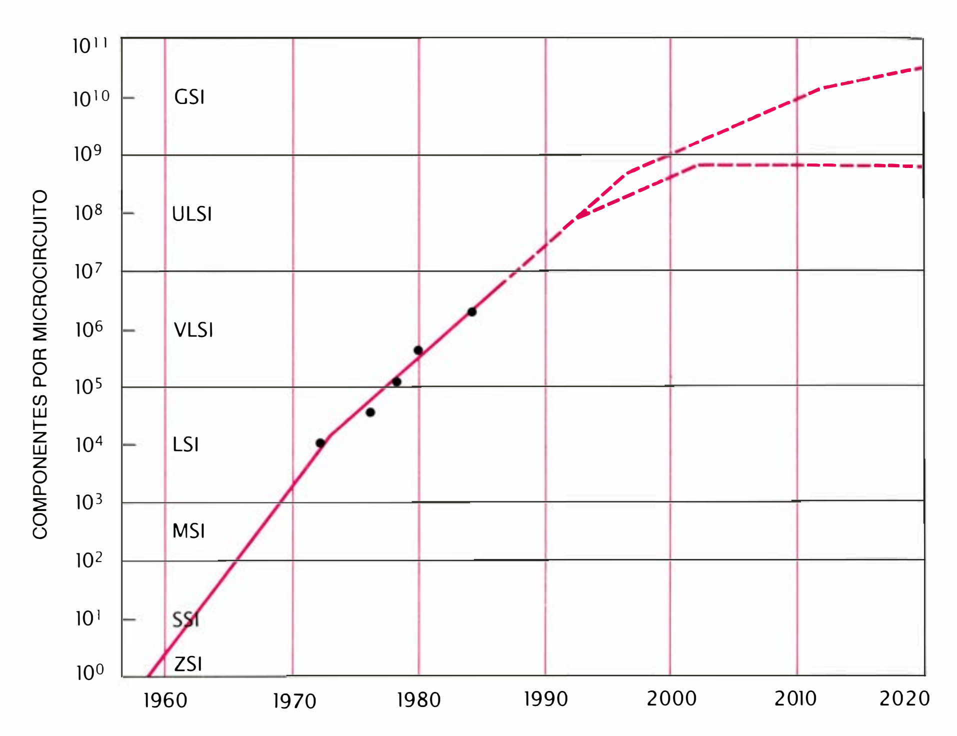




to relativamente bajo. Más todavía, su óxido natural, el dióxido de silicio (SiO_2), constituye un buen aislante que posee algunos atributos deseables para la construcción de circuitos integrados. Ningún otro material reúne semejante combinación de virtudes, lo cual contribuye a explicar la dominancia del silicio en la fabricación de circuitos integrados. No obstante, ciertas limitaciones intrínsecas a ese material han elevado en años recientes hasta un puesto de preeminencia a otro material semiconductor: el arseniuro de galio (GaAs).

La razón principal de que los diseñadores de microcircuitos recurran al arseniuro de galio tiene que ver con su velocidad. Al aplicársele un campo eléctrico débil, los electrones de la banda de conducción del arseniuro de galio se mueven seis veces más deprisa que los del silicio. Pero la movilidad de los electrones podría dar una idea exagerada de la superioridad del arseniuro de galio. Un criterio de comparación más preciso nos lo proporciona otra de las limitaciones intrínsecas del material, a saber, el tiempo que necesita un portador de carga, sometido a la influencia de un campo eléctrico de intensidad próxima a la de ruptura, para experimentar una caída del potencial eléctrico de un volt, por ejemplo. (Llámase campo eléctrico de ruptura al capaz de arrancar tantos electrones de valencia de los enlaces que definen el retículo cristalino que se provoca una avalancha autoionizante.) Si se utiliza esa medida, el arseniuro de galio es unas 2,5 veces más rápido que el silicio.

Así pues, en lo que atañe a ciertas limitaciones materiales, el arseniuro de galio aventaja al silicio en velocidad. Sin embargo, al reducirse el tamaño de los transistores, tal ventaja puede quedar desvirtuada por otros factores materiales. Puede lograrse que el transistor conmute más deprisa aplicándole un potencial superior, pero ello incrementa el calor a disipar por el dispositivo. En los ingenios extremadamente pequeños, la velocidad de conmutación puede hallarse limitada por la capacidad del sustrato para evacuar el calor generado. La conductividad térmica del silicio triplica la del arseniuro de galio; por ello, los diminutos dispositivos de silicio podrían cambiar de estado con la misma celeridad que los construidos con material ostensiblemente "más rápido".



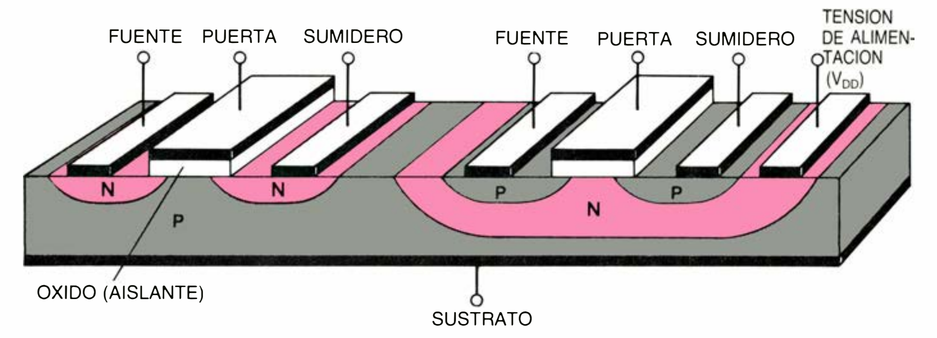
3. EL NUMERO DE COMPONENTES DE LOS MICROCIRCUITOS se duplicó anualmente durante el decenio de 1960. Hacia 1972, los diseñadores agotaron el espacio disponible para alojar componentes en la pastilla y aquel ritmo de crecimiento cedió. A pesar de ello, la estimación más optimista prevé que, hacia el año 2000, se habrá alcanzado la integración a gigaescala (GSI), es decir, la producción de un microcircuito dotado de mil millones de componentes. (Las estimaciones discrepan en los límites impuestos por los procesos de fabricación de microcircuitos: nos aproximamos al límite de resolución de los equipos de litografía óptica.)

Así, en cada nivel de la jerarquía es preciso equilibrar unas limitaciones con otras en mutua oposición. Ese proceso de compensación se aprecia con toda nitidez en el escalón de los dispositivos. Las limitaciones que les afectan son numerosas, pues en ellas participan todas las limitaciones de material, así como otras impuestas por el tamaño y la forma geométrica del dispositivo. Pese a tanta limitación, el problema fundamental para lograr la GSI en el nivel de dispositivo se acota con razonable claridad: determinar las mínimas dimensiones posibles de un transistor de efecto de campo de metal, óxido y silicio (MOSFET, por *metal-oxide-silicon field effect transistor*).

Un tipo corriente de MOSFET consta de dos islotes de silicio, dopado para incrementar la concentración de portadores de carga negativa (silicio de tipo *N*), insertos en una capa del mismo material, contaminado también para aumentar la concentración de portadores positivos (silicio *P*). A modo de puente entre ambos islotes, denominados fuente y sumidero, se tiende una capa de dióxido de silicio que sirve como aislante. Depositado sobre el aislante que va desde la fuente hasta el sumidero se dispone un electrodo metálico: la puerta. Cuando la puerta recibe una "entrada" positiva, se atraen electrones hacia la interfase del sustrato de silicio y el aislante. Los electrones forman un canal inducido, que permite el flujo de una corriente "de salida" desde la fuente hasta el sumidero, enviándose así una señal lógica a la siguiente etapa del circuito. En ausencia de señal de entrada no se forma canal entre los islotes, ni se produce corriente de salida.

Dado que la tecnología de MOS domina abrumadoramente en la electrónica digital, la consecución de la GSI depende de la progresiva reducción de escala del tamaño del MOSFET. El proceso de reducción de escala comienza por la definición de un factor de escala, a menudo denominado *S*. (Por ejemplo, si el factor de escala *S* es de 2, la altura y anchura del dispositivo se reducen cada una a la mitad de sus valores primitivos.) Además, la tensión de alimentación se reduce en el mismo factor. Así, la intensidad del campo eléctrico se mantiene fija y el disposi-

2. HISTORIA DE LOS MICROCIRCUITOS, que revela densidades de empaquetamiento cada vez mayores, según muestran estos ejemplos, fabricados entre 1959 y 1985 por la Fairchild Semiconductor Corporation. Arriba, a la izquierda (1959) aparece el primer transistor plano. A su derecha, el primer circuito integrado plano monopastilla (1961); consta de cuatro transistores y otros componentes. En el centro, a la izquierda, el circuito integrado lineal destinado a electrónica de consumo (1964), que porta cinco transistores. A su derecha una matriz lógica bipolar (1968) dotada de 180 transistores. En el ángulo inferior izquierdo (1978), el primer microcircuito capaz de contener, entera, una unidad central de procesamiento de 16 bits; tiene 20.000 transistores. Abajo, a la derecha, la CPU CLIPPER (1985), que aloja 132.000 transistores.



4. MOSFET COMPLEMENTARIOS, que constituyen la base de la tecnología de fabricación de circuitos integrados hoy predominante. Un MOSFET es un transistor de efecto de campo de metal, óxido y silicio. En un semiconductor como el silicio, la carga eléctrica pueden transportarla electrones (que llevan carga negativa) o "huecos" (portadores de carga positiva). En el silicio de tipo *N* predominan los electrones; en el de tipo *P*, los huecos. Cada tipo contiene, además, una pequeña cantidad de portadores de carga de signo contrario. Cuando la tensión de la puerta corresponde al estado "abierto", no fluye corriente desde la fuente al sumidero. Cuando la tensión corresponde a "cerrado", los portadores minoritarios son atraídos hacia la región situada bajo el electrodo de puerta, formándose un canal inducido que permite una gran corriente de salida del dispositivo. La ilustración muestra cómo se dispondrían sobre una única pastilla de silicio un par de MOSFET complementarios (uno cuyo canal fuera de silicio de tipo *N* y otro con canal de tipo *P*).

tivo no se ve sometido a esfuerzos superiores.

La reducción simultánea de tamaño e intensidad del campo eléctrico brinda notables ventajas. El tiempo necesario para la conmutación del dispositivo, que es función de la longitud del canal, disminuye en el factor *S*. La corriente disipada por unidad de superficie permanece constante, por lo cual no empeora el problema de evacuación del calor. La densidad de empaquetamiento de transistores en el microcircuito queda multiplicada por el factor S^2. Y, lo que quizás interese más, la energía consumida en cada operación de conmutación, que es función de la potencia y del tiempo de conmutación, queda dividida por el factor S^3. Consiguientemente, tras la reducción de la escala, el microcircuito consta de más dispositivos, que conmutan más rápidamente y que consumen menos corriente.

Ante tales ventajas, al diseñador de microcircuitos le gustaría saber hasta dónde puede llegar la reducción de escala. La respuesta depende de cuál sea la mínima longitud admisible para el canal del MOSFET, que viene dada, en parte, por la interacción entre la tensión de alimentación y la concentración del dopado, esto es, la concentración de átomos contaminantes (boro, de ordinario) del sustrato del canal, que es de tipo *P*. En cada unión entre materiales de tipo *N* y de tipo *P*, los portadores de carga emigran hacia el lado de la interfase que es rico en portadores de signo opuesto, por ser allí menor la concentración de portadores de signo propio; esto es: los electrones emigran hacia el lado de tipo *P* y, los huecos, hacia el de tipo *N*. De ese modo, en las inmediaciones de la unión se forma la región de carga espacial, desprovista de electrones y huecos. Para que funcione correctamente, el canal ha de tener al menos doble longitud que esa región.

Por supuesto, para hacer mínima la longitud del canal será deseable reducir la longitud de la región de carga espacial, lo que puede conseguirse aumentando la concentración del dopado. Sin embargo, la elevación de la concentración del dopado tiene el efecto, no deseado, de exigir un incremento de la tensión de alimentación. Mas, si ésta se eleva demasiado, el aislante de óxido situado bajo la puerta comenzará a rajarse por culpa del proceso de autoionización antes mencionado. Para lograr un canal de longitud mínima será preciso reconciliar la tensión de alimentación con la concentración del dopado. Se demuestra que, para los valores de tensión de alimentación y concentración del dopado de uso común, la longitud mínima teórica del canal oscila entre 0,1 y 0,2 micrometros (millonésimas de metro). De momento, tales límites son absolutamente teóricos; la longitud de canal de los MOSFET actuales oscila entre 1 y 2 micrometros.

Pero si la reducción de escala ofrece grandes ventajas para los dispositivos, ocurre lo contrario para el caso de las delgadas películas metálicas encargadas de interconectar los transistores de los microcircuitos. En primer lugar, no es cierto que todas las conexiones entre dispositivos se acorten al menguar el tamaño de los transistores. Durante los dos últimos decenios, mientras los transistores iban reduciéndose, las pastillas han ido creciendo. En consecuencia, las líneas de conexión de larga distancia (por ejemplo, las que enlazan transistores situados en rincones opuestos de la pastilla) han tendido a alargarse todavía más. Su aumento de longitud suele describirse mediante un "factor de escala de pastilla" (S_c), que guarda correspondencia con las dimensiones lineales del microcircuito.

Las conexiones locales, que enlazan transistores vecinos, sí se abrevian en proporción al factor de escala *S*. Pero ni siquiera en el caso de las conexiones locales resulta beneficiosa tal reducción. Dado el comportamiento eléctrico de los conductores, el tiempo requerido para trasladar una señal por una línea de interconexión quizá no decrezca al disminuir simultáneamente la longitud y la sección transversal del conductor; antes bien, el tiempo de retardo tiende a permanecer constante. Además, la densidad de corriente en el conductor crece con *S*. Si la densidad de corriente aumenta lo suficiente, el material del conductor puede resultar arrastrado por ella y dejar espacios vacíos. Tal fenómeno, denominado electromigración, acaba por arruinar la interconexión.

Avancemos un paso y examinemos las limitaciones en el nivel de circuito. En este escalón, uno de los principales problemas consiste en determinar la mínima tensión de alimentación a la que puede operar un circuito lógico. Tal vez el circuito lógico fundamental sea el inversor, que podemos hallar formando parte de la inmensa mayoría de la circuitería microelectrónica compleja. Un circuito inversor convierte una entrada de baja tensión (un "0" en el repertorio digital del ordenador) en una salida de alta tensión (un "1") y viceversa.

Y el dispositivo lógico más notable de la GSI quizá sea un inversor que consta de un MOSFET con un canal de tipo *N* y otro de canal de tipo *P*, combinación conocida por tecnología MOSFET complementaria. Salvo cuando se está procediendo a conmutar el dispositivo, sucede siempre que uno de los dispositivos se halla cerrado (en estado de conducción) y, el dispositivo complementario del par, abierto (en estado de no conducción). Por estar conectados en serie, el circuito apenas si consume energía mientras no se conmuta. En efecto, analizando las características del circuito se demuestra que la tensión de alimentación mínima necesaria para que funcione el inversor ronda, a temperatura ambiente, los 0,1 volt.

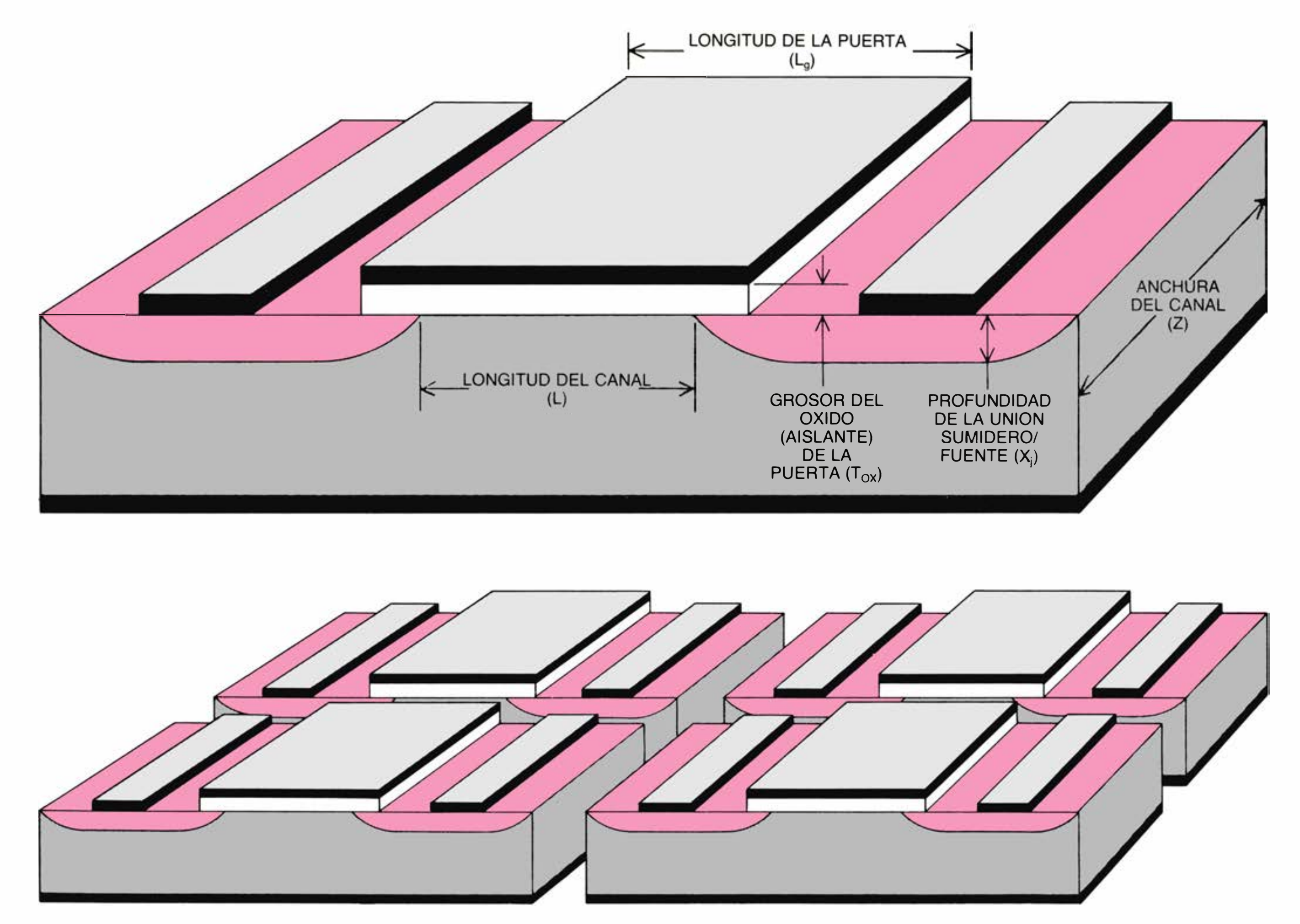
Una de las razones de la adopción inmediata de la tecnología MOSFET complementaria del procesamiento de información reside, precisamente, en su escaso consumo energético. Con ese cambio tiene que ver también otra limitación de la circuitería: la mínima longitud de su canal. Expuse más arriba que la longitud del canal condiciona

el nivel de dispositivo; ahora bien, las longitudes mínimas de canal correspondientes a los MOSFET de canal corto dependen también de la configuración de los circuitos donde se empleen. Se demuestra que una configuración MOSFET complementaria posibilita el trazado de canales más cortos, de 0,10 a 0,15 micrometros de longitud. El correspondiente incremento de velocidad que comporta el estrechamiento del canal es una de las razones que explican la situación privilegiada que hoy ocupa la tecnología MOSFET complementaria.

El peldaño más alto de la jerarquía corresponde al nivel de sistema; aquí el reto está en formular un modelo capaz de enlazar los parámetros de diseño de los niveles inferiores con la "arquitectura" global del sistema y su empaquetamiento. El único modelo de ese tipo del que tengo noticia lo formuló recientemente Brian Bakoglu, cuando era todavía uno de mis doctorandos de la Universidad de Stanford. El modelo, al que denominamos SUSPENS, se basa, en parte, en una fórmula conocida como regla de Rent; elaborada en el decenio de 1960 y ampliamente difundida entre los expertos en ciencias de cómputo durante el siguiente, dicha regla establece una relación entre el número de circuitos lógicos contenidos en una pastilla, por un lado, y, por otro, el número de conexiones necesarias para enlazar esos circuitos y el número de patillas de conexión necesarias para enlazar la pastilla con el resto del sistema. Una hábil generalización de la regla de Rent proporciona la longitud media de las interconexiones tendidas sobre la pastilla.

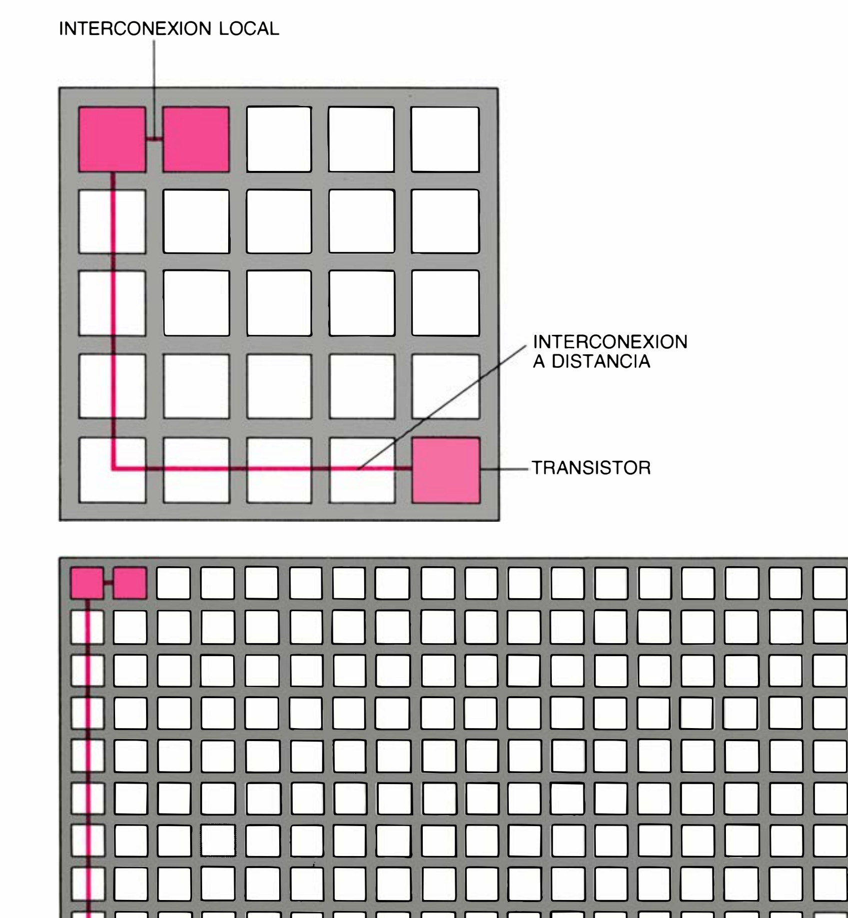
El modelo basado en la regla de Rent considera 26 parámetros. De ellos, ocho corresponden a medidas de la pastilla y de la arquitectura del sistema, según las presenta la regla propiamente dicha; cuatro pertenecen a la tecnología de construcción del transistor; cinco al conexionado interno de la pastilla, y nueve a conexiones entre módulos ("paquetes" del sistema, que contienen una o varias pastillas). Insertando en el modelo los valores correspondientes a esos 26 parámetros, se obtiene la frecuencia máxima del "reloj" que va marcándole el paso al sistema, así como el tamaño óptimo de la pastilla y la disipación total de corriente. Con esa información, el ingeniero determina el número óptimo de transistores por pastilla y ensaya el efecto que una nueva tecnología ejercería sobre el rendimiento del sistema.

Mediante la combinación de las condiciones impuestas en muchos niveles, el modelo del sistema resume la jerarquía entera desde la perspectiva teórica. Por otra parte, bastan tres parámetros (tamaño del rasgo mínimo, superficie utilizable de pastilla y eficacia de empaquetamiento) para describir la totalidad de los límites prácticos. El tamaño del rasgo mínimo constituye la menor dimensión lateral de los rasgos identificables en un MOSFET o en una interconexión metálica. Un ejemplo excelente es la longitud de la puerta de un MOSFET, algo mayor que la longitud del canal.



5. REDUCCION DE ESCALA, que ofrece considerables ventajas en el caso de los transistores MOSFET. Para reducir la escala de un MOSFET se empieza eligiendo un factor de reducción, *S*. Todas las dimensiones lineales se dividen por *S*, y lo mismo la tensión de alimentación. En este caso, *S* es 2. Como se advierte en el dibujo inferior, la densidad de empaquetamiento se multiplica por S^2. El retardo de puerta, que es el tiempo necesario para que un electrón o un hueco atraviese el canal, queda dividido por *S*. La densidad de disipación de corriente permanece constante, por haber quedado divididas por el factor de escala tanto la tensión como la intensidad que atraviesa los dispositivos. Y, lo que quizás interese más, la energía necesaria para la conmutación queda entonces dividida por S^3.

En 1960, el tamaño del rasgo mínimo de los circuitos integrados medía unos 25 micrometros. Hacia 1980, ese tamaño se había dividido por 10. Si el tamaño del rasgo mínimo pudiera seguir disminuyendo con ese ritmo, hacia el año 2000 sería de 0,25 micrometros, lo que no excede en mucho los límites de nivel de dispositivo y de circuito antes descritos. En razón de la proximidad de tales límites, parece improbable que el ritmo histórico de reducción del 10 por ciento anual pueda seguir manteniéndose a lo largo de la década de 1990. Tal conclusión se ve reforzada por el hecho siguiente: el rasgo mínimo ha comenzado a acercarse al límite de resolución de los equipos de litografía óptica utilizados para la fabricación de circuitos integrados (unos 0,5 micrometros), impuesto por las longitudes de onda más cortas de la radiación visible.



6. EL CAMBIO DE ESCALA no resulta tan beneficioso para las interconexiones entre dispositivos como lo es para éstos. Al tiempo que los dispositivos se reducían, las pastillas ("chips") aumentaban de tamaño. El factor de escala de la pastilla (S_c) mide el incremento proporcional de uno de los lados de la misma. La ilustración muestra cómo afecta el cambio de escala a la pastilla. Aquí, tanto S (el decremento proporcional de la dimensiones del dispositivo) como S_c son iguales a 2. Mientras que las interconexiones locales que enlazan dispositivos vecinos disminuyen con S, las interconexiones a distancia tienden a crecer con S_c.

Conforme iba menguando el tamaño del rasgo mínimo, aumentaba regularmente el tamaño medio de la pastilla utilizada en los microcircuitos. Las tendencias seguidas por la superficie de la pastilla suelen analizarse examinando su lado, esto es, la longitud del lado de una pastilla cuadrada que fuera de tamaño típico para su tiempo. En 1960, el lado de pastilla era de 1,4 milímetros; en 1980, de 8 milímetros. De continuar la misma progresión hasta el año 2000, se alcanzaría una pastilla de 50 milímetros de lado. Ciertas limitaciones de índole práctica sugieren, empero, que ese ritmo ha de aflojar. En litografía óptica, el campo de visión y la resolución son inversamente proporcionales. Dado que los circuitos se graban de una sola vez, y no por partes, la necesidad de resolver rasgos cada vez menores actúa como factor limitante del tamaño de la pastilla. Una predicción más verosímil sería, pues, que en el año 2000 el lado de la pastilla medirá entre 20 y 40 milímetros.

La eficacia del empaquetamiento mide, como su nombre indica, la densidad con que pueden empaquetarse los transistores en un microcircuito. Desde la década de 1960 ha aumentado espectacularmente, gracias a los avances registrados en los procesos de fabricación y a las innovaciones introducidas en circuitos y en dispositivos. Se creció con rapidez hasta el año 1972, cuando se les agotó a los ingenieros la superficie de pastilla no utilizada para transistores e interconexiones. A partir de entonces progresó a un ritmo más lento la eficacia del empaquetamiento, y los avances en la misma comenzaron a depender de ingeniosos diseños de los transistores, del aumento del número de pasos de enmascarado en los procesos fotolitográficos e incluso de la construcción de dispositivos tridimensionales, que se extendían por el anverso y reverso de la superficie de la pastilla.

El efecto combinado que han tenido los cambios en el tamaño de los rasgos mínimos, superficie de pastilla y eficacia de empaquetamiento ha sido el rápido crecimiento del número de componentes integrados en el microcircuito. Ese crecimiento constituye, en efecto, una de las tendencias tecnológicas más sobresalientes jamás observadas. El número de dispositivos por pastilla, que comenzó siendo de una unidad en 1959, fue duplicándose anualmente a partir de entonces. En el decenio de 1970 (a consecuencia de las limitaciones prácticas ya explicadas) esa tasa de crecimiento declinó, pero sólo hasta el

nivel todavía muy notable de cuadruplicación cada tres años. Esa tasa de incremento anual debería persistir hasta los primeros años del próximo decenio, durante el cual, los efectos de otros elementos limitadores tornen probablemente a frenar la tasa de crecimiento. A pesar de todo, la respuesta a la cuestión planteada al principio de este artículo es que la integración a gigaescala (GSI), es decir, de mil millones de componentes por microcircuito, debería alcanzarse antes del año 2000.

No obstante, los efectos de la GSI no se percibirán inmediatamente. La experiencia del pasado muestra que el volumen de producción de un nuevo producto microelectrónico no alcanza el máximo hasta tres o cinco años después de su presentación comercial. (Tal demora es necesaria para el crecimiento de la demanda y la plena capacidad de producción.) Por consiguiente, no es probable que los efectos definitivos de un microcircuito que contenga mil millones de componentes se dejen sentir antes de 2005. Así pues, ambas limitaciones, teóricas y prácticas, inducen a pensar que el ardoroso ritmo de la revolución informática se sostendrá, verosímilmente, durante los dos próximos decenios.

Atravesados los dominios de la teoría y la práctica, me apartaré ahora de la ciencia y la ingeniería para echar un vistazo a la historia, cuyas enseñanzas pueden servirnos para columbrar el futuro de la revolución informática. Los historiadores han observado con frecuencia que la evolución y el desarrollo de las técnicas comercialmente venturosas tienden a seguir una curva en forma de S, la llamada curva logística. Inicialmente el desarrollo es lento y se encuentra casi confinado al laboratorio. Aceptada la importancia industrial de la nueva tecnología, tras el éxito de los primeros tanteos comerciales, se le dedican sustanciosas inversiones; sigue un período de crecimiento rápido. En la cima de la curva se tropieza con limitaciones fundamentales, se frena el crecimiento y la tecnología en cuestión comienza a quedar obsoleta.

Son varios los signos indicadores de que la revolución informática está obedeciendo esta ley, como aconteció con la industrial que comenzó en el siglo XVIII. Trazando los paralelismos de esas dos revoluciones tecnológicas, quizá se localice en qué punto de la curva logística se encuentra la revolución informática. Nadie negará, supongo, que el material más importante de la Revolución Industrial fue el hierro. En ese sentido desempeñó un papel análogo al

PROPIEDAD	VALOR ANTES DEL CAMBIO DE ESCALA	VALOR TRAS EL CAMBIO DE ESCALA
LONGITUD LOCAL A GRAN DISTANCIA	 L L	 L/S S_cL
GROSOR	W	W/S
ANCHURA	Z	Z/S
RESISTENCIA LOCAL A GRAN DISTANCIA	 R R	 SR S_cS^2R
CAPACIDAD LOCAL A GRAN DISTANCIA	 C C	 C/S S_cC
TIEMPO DE RETARDO LOCAL A GRAN DISTANCIA	 RC RC	 RC $(SS_c)^2RC$
DENSIDAD DE CORRIENTE	J	SJ

7. ENTRE LOS EFECTOS de la reducción de escala se cuentan algunos inconvenientes serios. Al modificarse la escala de la red de interconexiones, los tiempos de retardo de los enlaces a larga distancia (el tiempo requerido por la señal para recorrer la conexión) tienden a crecer exponencialmente con el producto de S y de S_c. La densidad de corriente aumenta proporcionalmente a S, lo cual puede causar estragos en la interconexión y provocar la aparición de zonas vacías, que interrumpen por completo la comunicación.

del silicio, material clave de la revolución informática. La combinación del hierro con otros elementos, que produce distintas aleaciones conocidas por aceros, se corresponde con el dopado del silicio mediante átomos de impurezas para obtener el material de partida de los circuitos integrados. Es fácil continuar con la analogía en el nivel de dispositivo (el transistor y el pistón), de circuito (un circuito integrado y un motor de combustión interna) y de sistema (una red informática y un sistema de transporte).

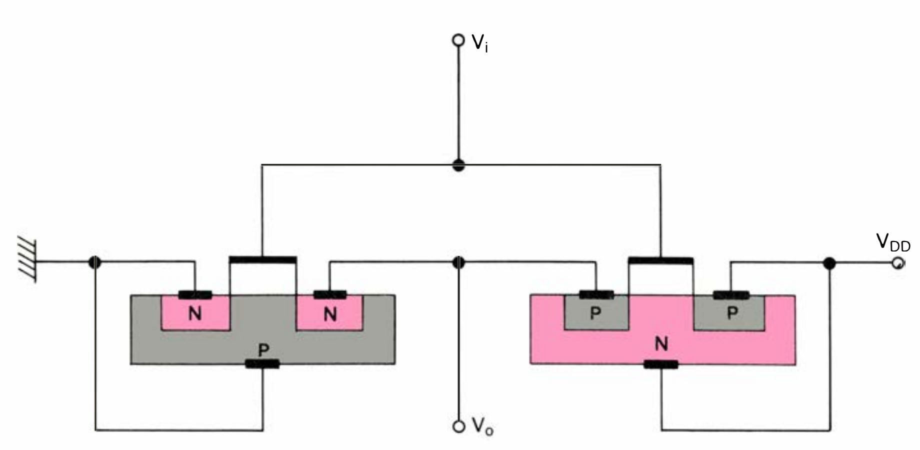
Ahora bien, de 1860 a 1900, la producción anual estadounidense de acero se duplicó cada cuatro años. Al final de ese período se alcanzó la rama superior de la S y se saturó la curva de producción. Desde 1900 hasta 1985 la producción de acero ha permanecido casi inalterable. La producción de silicio, por el contrario, se encuentra todavía en el tramo empinado de la S. Durante más de un decenio, la producción se ha estado duplicando cada dos años. Además, como he señalado ya, el logro de la GSI hace pensar que los rápidos incrementos persistirán durante un par de decenios más. Pero no mucho después se podrá llegar al hombro de la curva logística, estabilizándose la producción de silicio, o cuando menos, frenándose su rápido crecimiento.

Conviene señalar que, aun encontrándose en la rama superior de la S, el acero sigue constituyendo, por tonelaje, el material industrial dominante. Sin embargo, la producción de otros materiales, –aluminio, titanio, plásticos y materiales híbridos– crece mucho más deprisa que la de acero. Es verosímil que la industria de los circuitos integrados se ajuste a una pauta similar. El silicio será el material predominante, pero el centro de atención científica y de desarrollo económico se desplazará hacia otros materiales. Entre ellos se contarán los circuitos integrados basados en la superconducción criogénica, las finas películas de semiconductor depositadas sobre materiales aislantes y los sustratos formados por diversos materiales mediante deposición por haces moleculares. Tales posibilidades sugieren que el crecimiento no tiene por qué concluir cuando se alcance la rama superior de la curva.

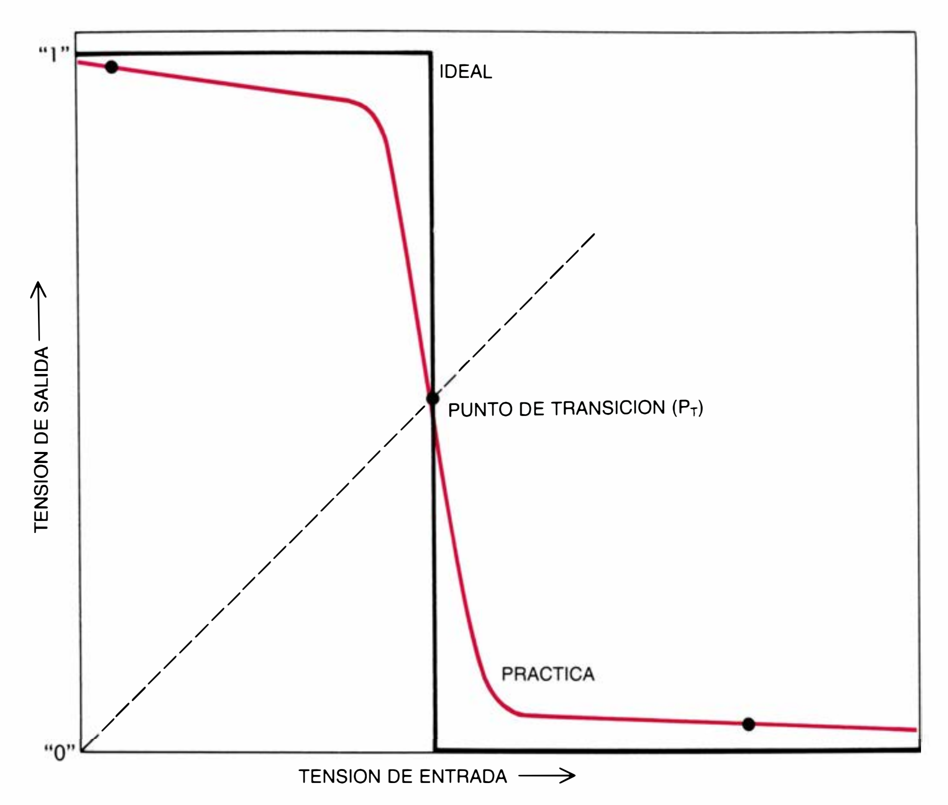
En efecto, un rasgo muy estimulante de la historia de la tecnología es la aparición de discontinuidades: puntos donde se rompen las limitaciones establecidas y se abren súbitamente nuevas perspectivas. Las discontinuidades responden a descubrimientos o invenciones que alteran de modo radical las expectativas futuras. Son varias las innovaciones que pudieran tener esos espectaculares efectos sobre la tecnología de los circuitos integrados. La mejor forma de comprenderlas consiste en examinar las escalas de longitud en las que manifestarán sus efectos.

Una escala nos la brinda la distancia que un electrón o un hueco pueden viajar en el seno del silicio antes de colisionar contra un átomo de silicio en vibración o un átomo contaminante. Se trata del recorrido libre medio; cuanto mayor sea esa distancia, tanto más deprisa pueden viajar los portadores. El recorrido libre medio puede aumentarse sin más que sustituir el silicio por arseniuro de galio, como se hace en el MESFET (esto es, *metal-semiconductor field-effect transistor*, transistor de efecto de campo de metal y semiconductor). Una propuesta más radical es el MODFET (*modulation-doped field-effect transistor*), en el cual se deposita una delgada película de arseniuro de aluminio y galio sobre un sustrato de arseniuro de galio que no contiene dopante. La ausencia de átomos de impureza aumenta el recorrido libre medio, lo que confiere a los MODFET más velocidad de operación que los MESFET, los cuales, a su vez, son más rápidos que los MOSFET.

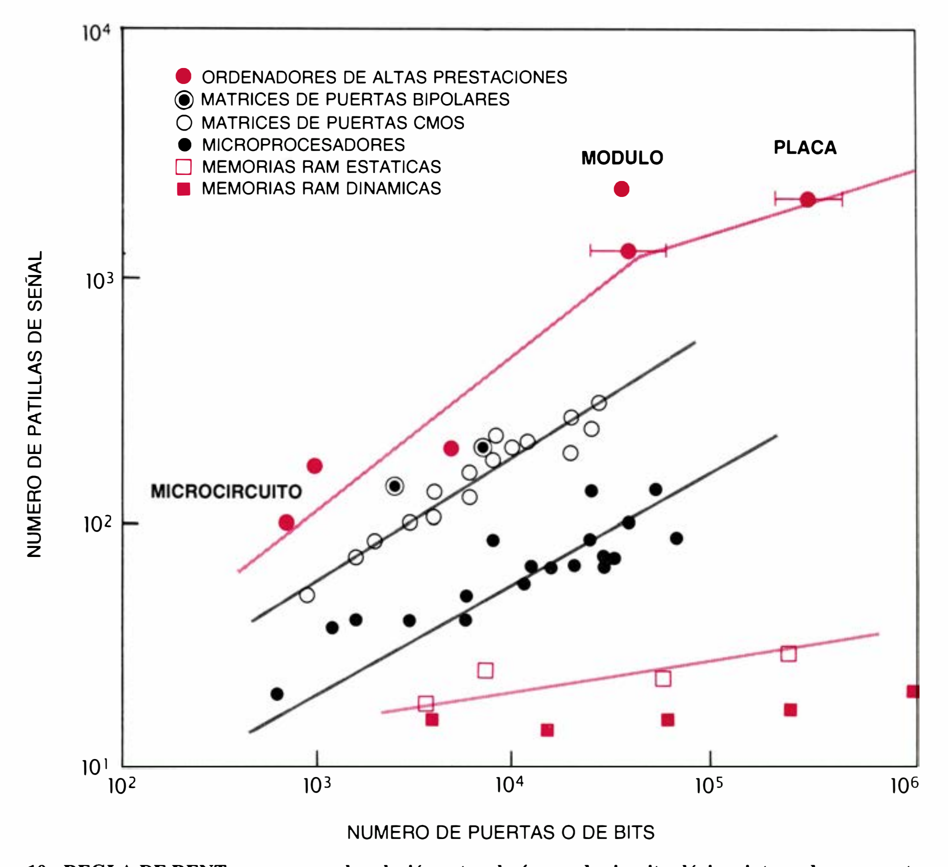
Puede incrementarse más la velocidad de los portadores en una escala justamente inferior a la del recorrido libre medio. Cuando la longitud del canal es



8. CIRCUITO INVERSOR, compuesto por dos MOSFET complementarios. El inversor es uno de los circuitos lógicos fundamentales de esa técnica. Convierte señales de entrada de tensión baja ("0") en tensiones de salida de valor alto ("1"), y viceversa. A menos que el circuito esté conmutándose, uno de los MOSFET permanece abierto (o sea, en estado de "no conducción") y el otro cerrado (en estado de conducción). Dado que los dispositivos están conectados en serie, el circuito consume poquísima potencia cuando no conmuta. Su escaso consumo energético constituye una notable ventaja de la tecnología MOSFET complementaria.



9. CURVA DE TRANSFERENCIA del circuito inversor que se representa en la figura 8, que muestra la tensión de salida en función de la tensión de entrada. Lo ideal es que la línea esté perfectamente cuantizada, es decir, que forme un escalón perfecto. De ser así, la pendiente de la línea en el punto de transición entre salida alta y salida baja resultaría ser infinita. Aunque en la práctica ello no resulta factible, la curva correspondiente al inversor complementario se aproxima al caso ideal más que la de cualquier otro circuito.



10. REGLA DE RENT, que expresa la relación entre el número de circuitos lógicos integrados que porta una pastilla y el de patillas de conexión necesarias para conectarla al resto del sistema. La regla es empírica, y se infirió durante la década de 1960 gracias a la mayor experiencia en el desarrollo de microcircuitos. La regla de Rent forma parte de los fundamentos de un modelo global que enlaza las propiedades de los materiales, de los dispositivos y de los circuitos con las características del sistema donde vayan a emplearse.

igual o menor que el recorrido libre medio de un electrón, resulta posible que un electrón atraviese el canal sin colisionar. Semejante desplazamiento sin colisión se denomina movimiento balístico. En distancias cortas, la velocidad balística puede ser varias veces mayor que la componente de velocidad de avance en una trayectoria de colisión, que es el tipo de trayectoria común en canales más largos que el recorrido libre medio. Los transistores balísticos constituyen otra frontera en la integración de circuitos, merecedora de detallada exploración.

En una escala todavía menor nos situamos en el régimen cuántico, que se extiende desde el recorrido libre medio, que en el silicio es de unos 10 nanómetros hacia abajo (un nanómetro son 10^{-9} metros). Una de las peculiaridades del MODFET antes mencionado es que consta de dos materiales distintos unidos en un único retículo cristalino llamado "superretículo". En virtud de las propiedades físicas de uno de los materiales (el arseniuro de galio), éste actúa a modo de "pozo cuántico" que captura muchos electrones del otro material (arseniuro de galio y aluminio). Con el concurso de la tecnología superreticular tal vez puedan trazarse en los cristales pozos cuánticos verticales y horizontales, cuya sección recordaría a un tablero de ajedrez. Los electrones que pasaran, por efecto túnel, de uno a otro escaque se aprovecharían para efectuar operaciones digitales. Los dispositivos de pozo cuántico decuplicarían la velocidad de los MOSFET del año 2000, por céleres que los imagine un cálculo optimista.

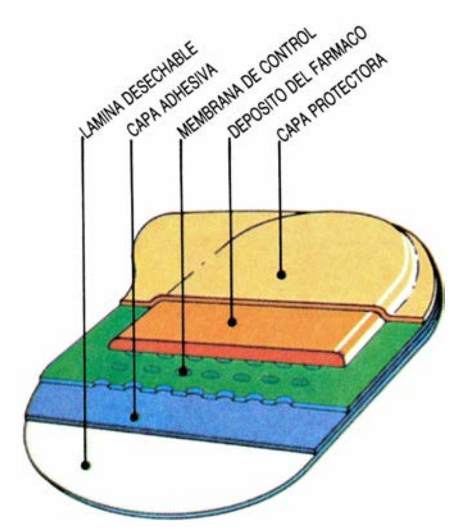
Todavía no se han obtenido dispositivos tridimensionales de pozo cuántico, pero su construcción sigue atrayéndonos desde el horizonte tecnológico. De resultar factible su fabricación, imprimirían a la curva logística un nuevo ascenso empinado. En realidad, la situación actual podría compararse con la del decenio de 1950, en el intervalo que medió entre la invención del transistor y del circuito integrado. Acabamos de citar el más rudimentario de los dispositivos de pozo cuántico, el MODFET. Habrá, pues, que ampliar la noción de pozo cuántico a estructuras tridimensionales, proporcionar conexiones entre ellas y situar tales dispositivos en el contexto de la nueva arquitectura de ordenadores. De lograrse todo ello, quizá se superen las limitaciones que pesan sobre el diseño y la fabricación de microcircuitos.

Ciencia y sociedad

Sistemas transcutáneos

Estos últimos años ha empezado a utilizarse en la práctica clínica un nuevo grupo de medicamentos: los sistemas terapéuticos transdérmicos (STT), dispositivos que permiten la administración cómoda e indolora del fármaco a través de la piel, en lugar de inyectado. El sistema puede compararse con la infusión sin aguja. La administración parenteral en la sangre, que elude la vía intestinal, evita las reacciones de degradación y toxicidad que entraña el primer paso por el hígado (el denominado "metabolismo del primer paso"), propio de la administración oral, así como la influencia de los diversos tiempos de tránsito intestinal, cambios del entorno y condiciones de absorción de las distintas secciones del tracto digestivo.

La piel sana y sin daños constituye una vasta envoltura impermeable; desempeña numerosas funciones, la principal de las cuales consiste en proteger al organismo de la deshidratación y las lesiones. Según se ha comprobado, por mero contacto con la piel pueden producirse intoxicaciones, o desencadenarse otros efectos indeseables, señal de que algunas substancias pueden atravesarla en cantidad suficiente. Los modernos métodos de bioanálisis han demostrado, además, que los medicamentos administrados a través de la piel no sólo ejercen su efecto localmente, sino que, en condiciones adecuadas, pasan a la circulación sanguínea en cantidades suficientes para emplearlos en terapias sistémicas.



1. Sistema transdérmico de difusión de medicamentos desarrollado por la compañía Ciba-Geigy. El fármaco contenido en el depósito se libera a ritmo constante del apósito y, penetrando en el organismo a través de la piel, pasa directamente a la circulación sanguínea; con ello se evitan los trastornos que puede provocar la administración por vía oral

La novedad que aportan los sistemas transdérmicos es que liberan el fármaco durante períodos más prolongados y a un ritmo fijado de antemano. Ello les distingue de las pomadas y ungüentos, con los que ya se venía aplicando productos de interés a través de la piel. En éstos, el control de la dosis es poco exacto, pues depende del grosor de la capa aplicada y de la superficie cubierta; además, actúan poco tiempo.

Externamente recuerdan un apósito, pero su estructura interna es bastante más compleja que la de éstos [*véase la figura 1*]. El elemento fundamental es un reservorio de medicamento, formado por una matriz polimérica maciza o sólo por una preparación líquida espesa limitada por una membrana. En la liberación del fármaco por difusión se distinguen, consiguientemente, los sistemas de control por matriz y por membrana. El depósito de medicamento está limitado, hacia el exterior, por una capa protectora impermeable; hacia la piel, la membrana de control se recubre con una capa adhesiva y, hasta la aplicación del sistema, además con una lámina desechable e igualmente impermeable. En los sistemas transdérmicos de control por membrana, el fármaco se libera con ritmo constante durante un lapso temporal calculado de antemano. La dosis que se administra al organismo se corresponde directamente con el tamaño del sistema.

La compañía Ciba-Geigy, convencida desde un principio de las ventajas que brinda para el paciente la terapia transcutánea, ha desarrollado, en colaboración con la norteamericana Alza Corporation, los primeros sistemas de ese tipo. Por ejemplo, el STT Scopoderm se emplea para combatir el mareo. El sistema, circular y de 2,5 centímetros cuadrados escasos, contiene escopolamina, de la que se liberan 0,5 milígramos a lo largo de tres días.

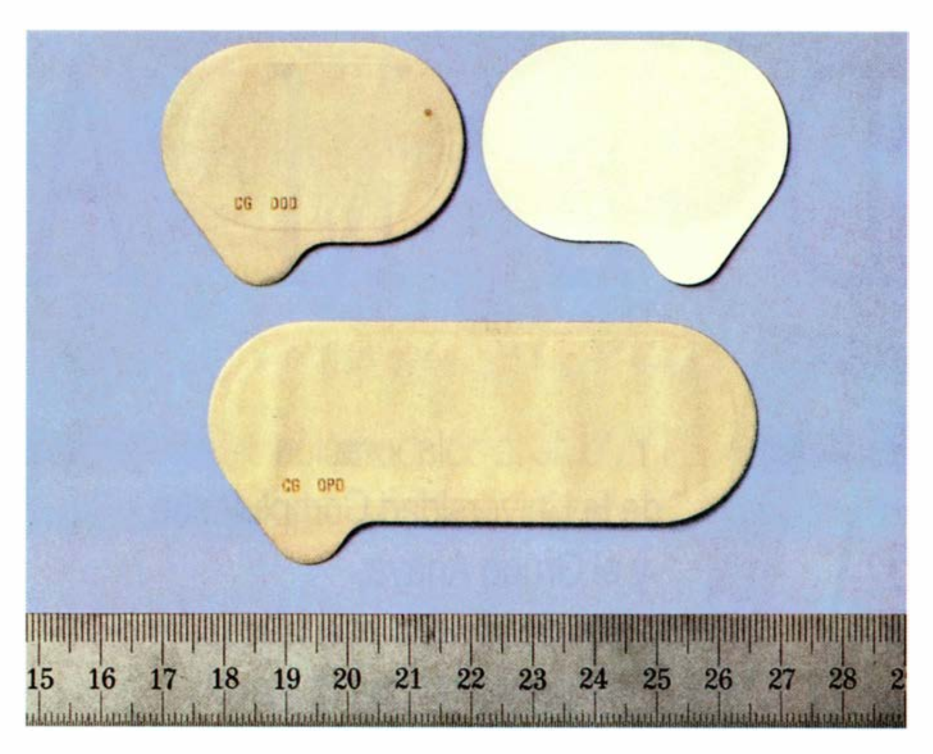
El aporte constante de cantidades muy pequeñas de medicamento elimina las náuseas sin provocar los molestos efectos secundarios de la escopolamina, asociados con tomas de concentración superior. No se conoce modo alguno de administrar ese medicamento que alcance resultados terapéuticos comparables con los de la infusión constante en la sangre. El STT Scopoderm se sitúa detrás de la oreja; en ese punto, la permeabilidad de la piel a la escopolamina duplica la de cualquier otro del tronco superior. Se suspende la administración del fármaco sin mayores problemas retirando el sistema.

El sistema contra el mareo no se utilizará, en general, de modo habitual o durante largos períodos; otros sistemas transdérmicos, por el contrario, se han ideado precisamente para un uso prolongado, de años incluso. Vale de ejemplo el STT Nitroderm, cuyo compuesto activo es la nitroglicerina [*véase la figura 2*]. Desde hace más de un siglo, y hasta la fecha, se emplea esa substancia con fines terapéuticos, en particular en la lucha contra la angina de pecho: ensancha los vasos que irrigan el corazón, estrechados durante los episodios espasmódicos.

Entre sus desventajas se cuenta el que la acción de la nitroglicerina resulta fugaz: el organismo la degrada con velocidad inusitada. Por ello, para cortar los ataques, el fármaco suele tomarse en forma de tabletas, que se mastican. Incluso en las preparaciones denominadas "retardantes", que liberan el medicamento lentamente, se limita su acción a una pocas horas.

La aplicación del sistema transdérmico de difusión continua modificará por fin esa situación: permitirá la administración continua y preventiva de nitroglicerina durante 24 horas. El enfermo queda protegido de los dolorosos ataques sin más que aplicarle diariamente un apósito en el tronco superior o el antebrazo. Para adecuar la dosis de medicamento al paciente se dispone de sistemas de varios tamaños. El método ha venido a brindar una nueva modalidad de administración de ese fármaco, de virtudes sobradamente conocidas, a la vez que mejora la calidad de vida de los pacientes, obligados a tomar nitratos de por vida.

El último sistema transdérmico desarrollado por Ciba-Geigy, el STT Estraderm, diseñado también para aplicaciones prolongadas, sirve la hormona estradiol. Se empleará para remediar las molestias que provoca en las mujeres el déficit de estrógenos durante los años de transición en los que, tras el descenso de su síntesis, ésta cesa por completo. El componente activo pasa directamente del dispositivo de aplicación a la sangre, evitándose su degradación hepática y, por consiguiente, bastando una fracción de la dosis oral para alcanzar la sustitución fisiológica de la hormona. Los niveles hormonales se incrementan hasta los valores ante-



2. Ejemplares del sistema terapéutico transdérmico "Nitroderm", desarrollado para prevenir la angina de pecho. Se ofrecen varios tamaños, para adecuar la dosis administrada a las necesidades del paciente

riores al descenso, sin cargar el organismo con los productos de degradación del fármaco que se formarían en el hígado. El STT Estraderm es un apósito transparente, circular u ovalado, que se cambia dos veces por semana y que puede aplicarse en varias partes del cuerpo, preferiblemente en las nalgas o el muslo. El médico puede optar por tres tamaños de dispositivo, en función de las necesidades de la paciente.

Los sistemas transdérmicos pueden considerarse dispositivos terapéuticos especialmente adecuados a los pacientes, puesto que son de aplicación externa y no causan dolor, y porque el tratamiento puede suspenderse sin mayores contratiempos. También brindan al médico ventajas respecto de otras formas anteriores: se reducen los efectos colaterales, se prolonga la actividad de substancias que se degradan y se elude el tracto digestivo sin recurrir a la inyección con jeringuilla. No sorprende, por consiguiente, que, ante tan halagüeñas perspectivas, se haya emprendido el desarrollo y ensayo de sistemas transdérmicos portadores de otros principios activos, o se ofrezcan ya en el mercado. Por supuesto, no todas las substancias se adecúan a la aplicación transdérmica, pero hoy centran el interés por su posible aplicación en esos sistemas los preparados con fármacos contra la hipertensión arterial y otras enfermedades cardiovasculares, así como analgésicos, antiasmáticos, etcétera. Entre otras condiciones, habrá que comprobar si, manteniendo un tamaño adecuado del sistema, la cantidad de medicamento que atraviesa la piel es suficiente y si el tratamiento o sustitución de substancias sintetizadas por el organismo se prolonga durante intervalos temporales significativos. (Henning Asche.)

Lectores miopes

¿Puede un exceso de lectura provocar miopía? Así parece deducirse de una larga historia de observaciones, incluido un informe que recoge el aumento de los casos de miopía entre la población esquimal a partir de la implantación de la enseñanza obligatoria en esa comunidad. Las investigaciones también sugieren una asociación de la miopía con el alargamiento del ojo: la luz procedente de imágenes distantes tiende a enfocarse por delante de la retina, en vez de sobre ella. ¿Cuál es, sin embargo, el vínculo entre la lectura y el alargamiento del globo? El equipo dirigido por Josh Wallman, del City College de la Universidad de la ciudad de Nueva York, ha propuesto una hipótesis fundamentada en experimentos realizados recientemente con pollos. Si se subestimulan las regiones periféricas de la retina, informan esos especialistas en *Science*, la reducción de la actividad neuronal en esas zonas podría provocar el alargamiento de ciertas partes del globo, tornándose miope el ojo entero.

Para verificar su hipótesis, el equipo tapó uno de los ojos a pollos recién nacidos (el ojo libre servía como control). A un grupo se le cubrió la totalidad del ojo con un plástico translúcido (imagínese mirando a través de una pelota de ping-pong: se ve la luz pero no se distinguen las formas). A otro grupo sólo se le tapó parte del ojo, y por tanto sólo parte de la retina.

Las mediciones efectuadas a las dos y a las seis semanas de edad evidenciaron que todos los ojos tapados se habían alargado y padecían miopía. Pero se advirtieron también diferencias importantes entre los grupos. En los pollos a los que se les había tapado el ojo de modo parcial, el alargamiento sólo se había producido en las zonas desprovistas de visión. Los ojos que se habían tapado completamente mostraban, en cambio, un crecimiento uniforme. Wallman sugiere, en consecuencia, que la miopía quizá sea fruto del control que ejerce la retina sobre el crecimiento local del ojo.

¿Por qué esa importancia del control local? Al nacer, los humanos y los pollos son hipermétropes, es decir, ven bien de lejos: el globo es demasiado corto respecto de la potencia óptica del cristalino y los rayos lumínicos convergen por detrás de la retina. En el transcurso del desarrollo normal, tiende a incrementarse la longitud del ojo, aproximándose el plano de la imagen a la retina. Según la teoría de Wallman, cada región retiniana ejerce un control independiente sobre el crecimiento. Puede que al registrarse una intensa actividad neuronal en una zona (es decir, cuando se ha alcanzado la agudeza visual) el crecimiento se frene por acción de inhibidores. Y a la inversa: una actividad escasa estimularía la secreción de promotores del crecimiento.

Wallman propone que en la lectura no se estimulan las neuronas situadas en la periferia de la retina. Afirma el autor que las neuronas del centro retiniano, la fóvea, se fijan en porciones muy pequeñas del campo visual, como, por ejemplo, el punto de la *i*. Las variaciones entre letras y de la posición del espacio en blanco a lo largo de un página brindaría suficiente estímulo para que se disparase en esa región gran número de impulsos nerviosos. Por el contrario, las neuronas de la periferia cubren campos visuales cuando menos diez veces mayores; el rasgo en negro se promedia con el fondo blanco y la neurona vislumbra un gris uniforme: se consumen las respuestas de las neuronas que no reciben estímulo y esas zonas del globo se alargan y provocan la miopía del ojo entero.



Programas para computación avanzada

Las máquinas de computación en paralelo plantean un reto fundamental: redactar programas que realicen muchas cosas simultáneamente. ¿Cómo organizar una colmena de actividades para que formen un todo coherente?

David Gelernter

Los ordenadores son demasiado lentos. Ni siquiera la máquina convencional más veloz puede habérselas con más de un par de centenares de millones de multiplicaciones por segundo. ¿Qué hacer entonces con los problemas difíciles, por ejemplo, la simulación de las interacciones en sistemas de átomos? Incluso en los sistemas atómicos, tales simulaciones pueden requerir semanas de funcionamiento continuo de las máquinas más rápidas hoy disponibles. Son muchos los problemas importantes cuya dificultad es por lo menos igual a la de los aludidos: desde la simulación de la mutua atracción gravitatoria de objetos celestes, pasando por problemas de mecánica de fluidos, que desempeñan un papel fundamental en la predicción meteorológica, hasta el examen rápido de ingentes bases de conocimientos, cuestión esta que adquiere relevancia creciente en inteligencia artificial.

Una solución provisional consiste en tratar de conferir más velocidad al ordenador secuencial al uso, que ejecuta las instrucciones una tras otra por medio de una unidad central de procesamiento, mejorando sus circuitos y su organización interna. Pero la solución a largo plazo reside en el paralelismo, esto es, en interconectar muchos ordenadores y aplicar toda su potencia de cómputo conjunta a la resolución de un mismo problema. Diez ordenadores idénticos, que trabajasen en tándem, resolverían un problema, en el caso óptimo, diez veces antes que un ordenador que operara en solitario. Esa es la razón de que se encuentren ya en el mercado ordenadores en paralelo que disponen de decenas, centenares o millares de procesadores, de subordenadores. La fabricación de una computadora de funcionamiento en paralelo plantea varios problemas técnicos, muchos de ellos relativos a las vías de comunicación entre los procesadores [véase "Nuevas arquitecturas para computación avanzada", por Geoffrey C. Fox y Paul C. Messina, en este número]. A diferencia de las estrategias para dotar de mayor velocidad a los ordenadores tradicionales, que sólo obligan a ligeros retoques en el estilo de programación, la computación en paralelo plantea también serios problemas de programación.

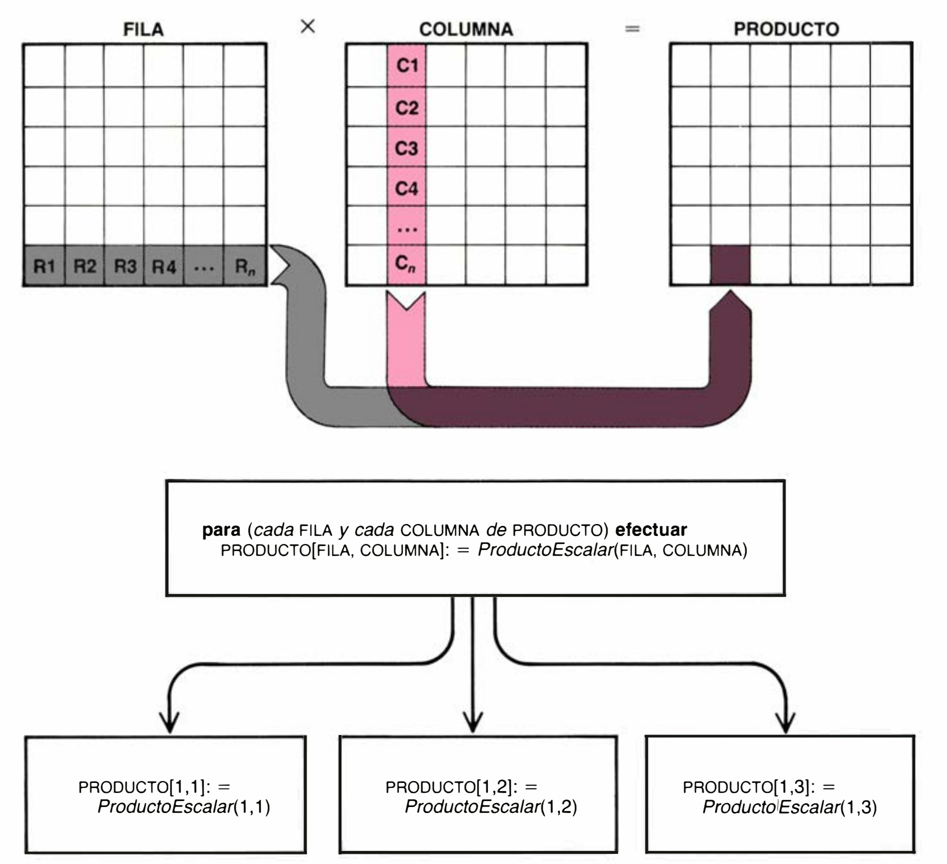
Representémonos un programa de ordenador como una suerte de máquina. Los programas para computación en paralelo son máquinas que se diferencian de los programas secuenciales tanto como el ordenador en paralelo se distingue del secuencial. Cierto es que el programa constituye un documento, un conjunto de instrucciones que ha de ejecutar el ordenador, redactadas en un lenguaje que entiendan tanto el programador como el ordenador. Sin embargo, al ser ejecutadas sus instrucciones, el programa se convierte en un acontecimiento, es decir, en un proceso que transforma datos en resultados. Ambos aspectos del programa convergen en el concepto de máquina. El texto es la máquina antes de ponerla en marcha; el acontecimiento es la máquina en funcionamiento.

Todo cuanto sea capaz de efectuar un programa puede llevarlo a cabo un "ingenio computador" mecánico, hecho de engranajes y ruedas dentadas. Empero, resulta mucho más fácil fabricar la máquina-programa, pues su construcción sólo requiere una terminal de ordenador y un lenguaje de programación. Siendo el propio programa una máquina, ¿cuál es papel que desempeña el ordenador propiamente dicho, la máquina tangible en la cual funciona el programa? Se convierte en una suerte de fuente de energía para la máquina-programa. Ahí reside la causa de que el mismo programa pueda funcionar sobre una amplia variedad de ordenadores dispares y de que las características del programa sean bastante independientes de los pormenores físicos del ordenador que lo aloja.

Podemos, pues, examinar el procesamiento en paralelo centrando nuestra atención en los programas y echando sólo algún vistazo a la maquinaria tangible cuando programación y soporte físico se hallen entretejidos. El programa tradicional efectúa una sola cosa por vez, mientras que el paralelo crea y gestiona muchas líneas de actividad simultáneas. Tiene que descomponer las tareas en muchos fragmentos y coordinar su operación estableciendo un flujo organizado de instrucciones, datos y resultados.

Para construir un tipo nuevo de máquina-programa precisamos un juego completo de piezas componentes de un nuevo tipo, esto es, un nuevo lenguaje de programación. Los lenguajes de programación tradicionales carecen de la sintaxis y el vocabulario requeridos para especificar el paralelismo. Descritos burdamente, le permiten al programador dictar órdenes del siguiente tenor: "Haz *A*, luego *B* y después *C*"; pero no "Haz *A*, *B* y *C* al mismo tiempo". Ahora bien, esta necesidad prác-

1. LAS ABEJAS EJECUTAN un programa en paralelo cuando cuidan la colmena. Lo mismo que las abejas (agentes débiles que trabajan concertadamente), el programa organizado en paralelo puede aplicar una gran potencia de cómputo sobre un problema creando varios procesos, o *loci*, de actividad. Esos artrópodos coordinan sus actividades merced a señales visuales y químicas; análogamente, los procesos que se desarrollan simultáneamente en un programa en paralelo han de comunicarse entre sí, para operar en conjunción.



2. MULTIPLICACION DE DOS MATRICES, ejemplo típico de los problemas en los que un programa secuencial ordinario puede transformarse automáticamente en programa en paralelo. Los elementos de la matriz producto se obtienen multiplicando "escalarmente" las filas de la primera matriz por las columnas de la segunda, es decir, multiplicando ordenadamente cada elemento de una fila por el que ocupa la misma posición en la columna, y sumando todos los productos obtenidos (*arriba*). Los programas convencionales de multiplicación de matrices se centran en una subrutina de multiplicación escalar encargada de calcular los productos de los pares de elementos, alojada en el interior de un bucle que calcula los sucesivos productos escalares (*centro*). Dado que cada producto escalar puede calcularse con independencia de los otros, un compilador de paralelización convertiría automáticamente el bucle en una serie de actividades simultáneas.

tica de un nuevo lenguaje no es sino manifestación superficial de una cuestión más profunda. Los lenguajes de programación no adquieren existencia por sí mismos; son expresión de un modelo de programación determinado, de cierta manera de abordar la construcción de programas. Por consiguiente, la cuestión fundamental pasa a ser "¿Qué aspecto debe tener un programa de computación en paralelo?" Aunque son muchas las líneas de investigación que se siguen, hay tres grandes tipos de respuestas a la pregunta anterior.

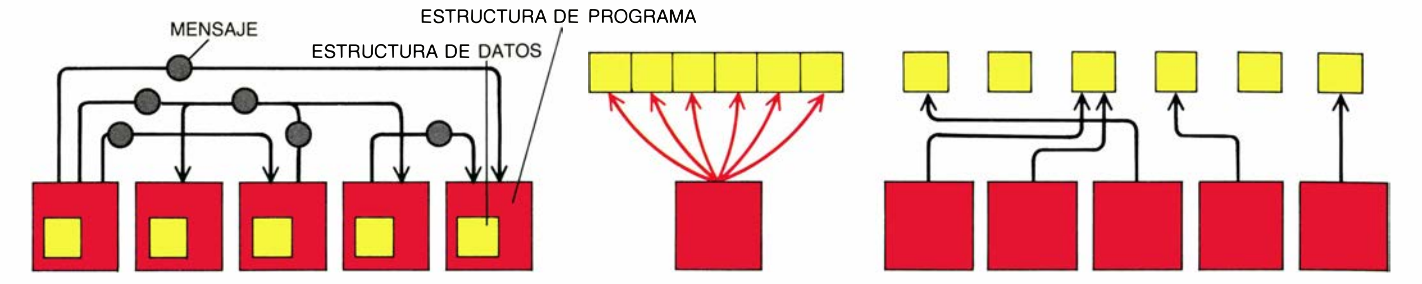
La primera escuela deja de lado la cuestión. No propone nuevos modelos de programación, sino que descansa en la generación automática de versiones en paralelo de los programas en uso. Una de las estrategias utilizadas en esta escuela es generalización de la que permite ejecutar un mismo programa en muchos ordenadores distintos. El ordenador suele necesitar ayuda para comprender un lenguaje de programación. En lenguajes "de alto nivel", como Pascal o Fortran, basta una línea para asignar a una variable el resultado de multiplicar dos números; pero la ejecución de tal orden comporta cierto número de operaciones en el seno de la máquina, que dependen del diseño de cada ordenador. Así pues, la orden ha de traducirse en la serie correcta de instrucciones a realizar paso a paso en código de máquina, que es el lenguaje digital primitivo propio del ordenador. De esa traducción se encarga un segundo programa, el llamado compilador.

Movidos fundamentalmente por el famoso "problema del mazo polvoriento", diversos grupos de trabajo de la Universidad de Illinois, la Universidad Rice, Multiflow Computer, Inc., IBM Corporation y de otras entidades, han puesto a punto compiladores que transforman automáticamente programas secuenciales en código de máquina paralelo. (Se llama "mazo polvoriento" la colosal inversión de tiempo y dinero que ha supuesto la confección de los programas ya existentes; el "mazo polvoriento" era, en realidad, un viejo paquete de tarjetas perforadas.) Los compiladores de paralelización inspeccionan el texto de los programas secuenciales en busca de operaciones que puedan ejecutarse a la vez, aunque el programa secuencial especifique que deben llevarse a cabo una tras otra. El compilador se encarga entonces de generar un código que refleje el paralelismo implícito que ha descubierto.

Por ejemplo, muchos programas se centran en un bucle computacional donde se realiza una misma operación, o idéntica serie de operaciones, con cada uno de los valores de una sucesión. Suele ocurrir que un compilador paralelizante es capaz de determinar, por inspección, si las iteraciones del bucle tienen forzosamente que realizarse en sucesión, según prescribe el programa. Cuando cada nuevo cálculo es función de los valores ya calculados en una iteración anterior, se requiere la secuenciación. De no ser así, el compilador podría convertir el bucle en código de máquina en paralelo, y posibilitar la realización simultánea de todos los cálculos.

Sea un programa para multiplicar dos matrices, un procedimiento fundamental del cálculo numérico. Cada elemento de la matriz producto representa el "producto escalar" de una fila de la primera matriz por una columna de la segunda. Para obtener el elemento situado en la fila i, columna j, de la matriz producto, se multiplican ordenadamente los elementos de la fila $i$ésima de la primera matriz por los elementos correspondientes de la columna $j$ésima de la matriz, es decir, primer elemento de la fila i por primer elemento de la columna j, segundo elemento de la fila por segundo de la columna, y así sucesivamente. Un programa convencional de multiplicación de dos matrices constará de dos bucles, encajado uno en otro; el interno, para la multiplicación de elemento por elemento, y el externo, para calcular los sucesivos productos internos. En cada iteración, el bucle exterior toma una fila y una columna, multiplica mediante el bucle interno y suma los productos.

Por ser autónoma cada iteración, desde un punto de vista lógico, el compilador paralelizante convertiría el bucle en código paralelo. Si multiplicáramos dos matrices cuadradas de orden 500 cada una, al paralelizar el bucle exterior y realizar simultáneamente los cómputos correspondientes a los productos interiores, en el caso óptimo el cómputo se aceleraría 250.000 veces. ("En el caso óptimo" significa aquí "en un ordenador que contiene 250.000



3. A LOS ESTILOS DE COMUNICACION compete definir las diferencias entre los distintos modelos de programación en paralelo. El modelo de transmisión de mensajes establece muchos subprogramas (*izquierda*), cada uno con sus propias estructuras de datos (tablas, listas y otras formas de organizar los datos, a los que las estructuras activas de los subprogramas se encargan de transformar). Los subprogramas tienen vedado utilizar las estructuras de datos de los demás; sólo pueden comunicarse enviando mensajes. Otra técnica, utilizada en la "máquina de conexión", consiste en que muchas estructuras de programa idénticas ejecuten una instrucción emitida por un programa "jefe", para transformar simultánea y sincrónicamente muchas partes de una estructura de datos (*centro*). Una técnica emparentada con ésta consiente que estructuras de programa no sincronizadas operen a la vez sobre una estructura de datos (*derecha*).

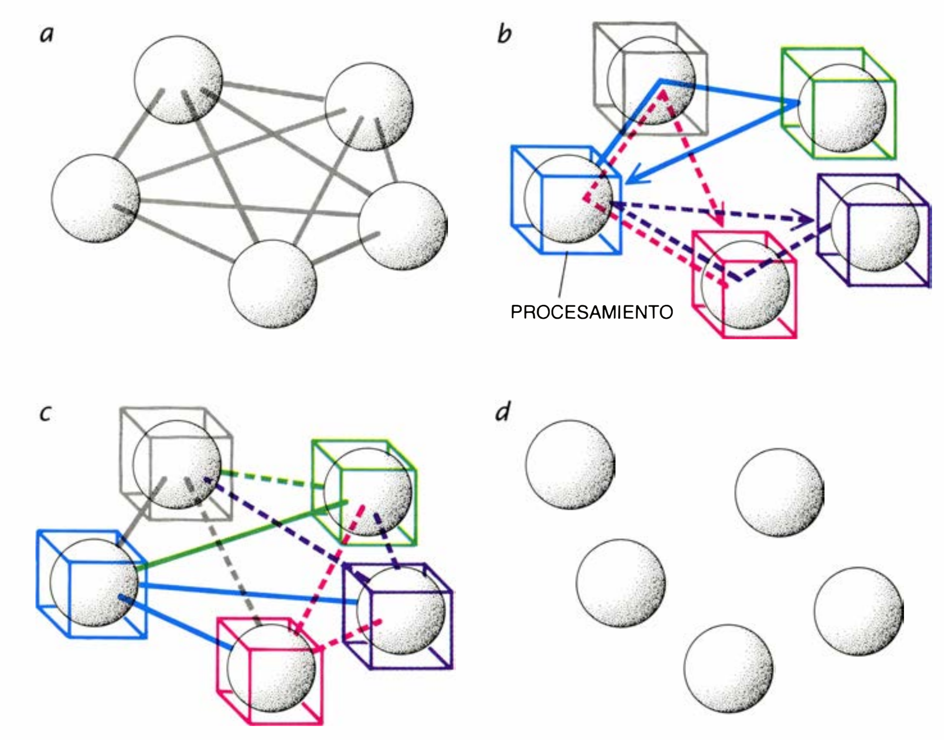
procesadores". No existen máquinas de ese tamaño de uso general; el número de procesadores de las máquinas disponibles, mucho menor, impone un límite superior a la aceleración del cálculo, pues fuerza a cada procesador a dar secuencialmente varios de los pasos que podrían ejecutarse en paralelo.)

Los ordenadores de flujo de datos ofrecen otro método de aprovechar el paralelismo implícito. No son máquinas ideadas para abordar el problema del mazo polvoriento, pues suelen diseñarse en conjunción con un nuevo estilo de programación. Pero se plantean la misma cuestión básica: cómo evitar modelos de programación explícitamente paralelos. Los programas de las máquinas de flujo de datos no especifican una secuencia de ejecución. Las únicas restricciones que pesan sobre su orden de ejecución son las que vienen impuestas por la estructura lógica del programa. Pueden ejecutarse simultáneamente dos instrucciones cualesquiera mediante sendos procesadores, a menos que una de las instrucciones necesite un valor de cuyo cálculo sea la otra responsable directa o indirecta. En tal caso, la ejecución se pospone hasta la llegada del dato requerido. Conforme va ejecutándose el programa, los datos fluyen de una instrucción a otra.

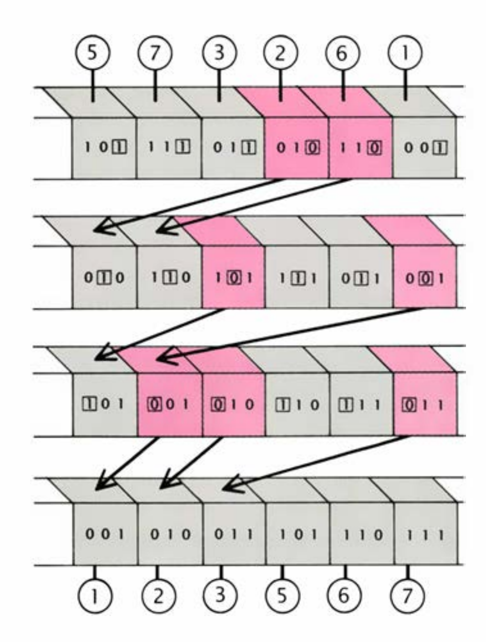
Así, en cierto sentido, es el propio ordenador y no el programa quien establece líneas de actividad paralela. Los lenguajes habitualmente asociados con las máquinas de flujo, los llamados "lenguajes funcionales", facilitan a la máquina la descomposición de la tarea en actividades paralelas. Esos lenguajes simplifican las relaciones de datos entre los enunciados y consienten al programador definir el valor de una variable solamente una vez. Establecido el valor de la variable, otras partes del programa pueden utilizarla libre y simultáneamente, sin preocuparse de cuándo y cómo cambiará su valor.

Los trabajos en curso sobre máquinas de flujo de datos y sobre lenguajes funcionales (que se están desarrollando en el Instituto de Tecnología de Massachusetts, la Universidad de Manchester y la Universidad de Utah, entre otras instituciones) iluminarán la operatividad de ese intento de abordar automáticamente el paralelismo. Es preciso señalar, igualmente, que los lenguajes funcionales no están estrictamente ligados a las máquinas de flujo de datos. Algunos investigadores consideran que su simplicidad y cercanía a los modelos matemáticos de programación los hace interesantes por derecho propio, con independencia del paralelismo. Otros (Paul Hudak, de la Universidad de Yale, constituye un ejemplo notorio) se están sirviendo de ellos para proporcionar bases claras y elegantes sobre las que construir lenguajes paralelos, a los que nos referiremos seguidamente.

Pudiéndose determinar automáticamente el paralelismo, ¿por qué ir más lejos? ¿A qué molestarse en crear modelos de programación nuevos y paralelos? La verdad es que las transformaciones automáticas sólo funcionan bien para programas de ciertas clases.



4. PROGRAMA DE TRANSMISION DE MENSAJES que permite computar los movimientos relativos de varios objetos que interactúan gravitatoriamente (*a*). Cada objeto tiene asignado un proceso, o subprograma. En cada ciclo de cómputo, cada proceso hace circular un mensaje que comunica la posición del objeto a su cargo a la mitad de los demás procesos, y es visitado a su vez por mensajes similares procedentes de la mitad de los procesos restantes (*b*). Al recibir un mensaje, los subprogramas calculan la fuerza entre su objeto y el objeto del que procede el mensaje en cuestión y añaden el resultado al propio mensaje que emiten ellos. Así, al regreso del mensaje, los procesos son informados de la mitad de las fuerzas que actúan sobre su objeto y calculan por sí mismos la otra mitad (*c*). Seguidamente, actualizan la posición de su objeto (*d*).



5. PROGRAMA para la máquina de conexión que permite ordenar de menor a mayor una sucesión de números, cada uno de ellos asignado a un único proceso. Todos los procesos examinan a la vez la última posición (el último dígito binario significativo) de la expresión binaria del número. Los que tienen un 0 se desplazan hacia la cabeza de la sucesión: se modifica su posición entre los elementos de memoria. Se repite el procedimiento para cada cifra binaria; los números que tengan un mismo dígito binario en la posición que se esté examinando conservan sus posiciones relativas. Tras un número de pasos igual al número de dígitos del máximo de los valores a ordenar, la serie queda ordenada de menor a mayor.

Los compiladores paralelizantes, por ejemplo, reescriben los programas secuenciales en código de máquina paralelo basándose en el comportamiento esperado del programa. Muchos programas —los de cálculo numérico, en particular— tienen un desarrollo predictible y, por tal razón, el compilador puede determinar fiablemente qué partes de la tarea resultan viables en paralelo. No obstante, las pautas de ejecución de algunos programas son complicadas o impredecibles, y cuesta entonces descubrir el paralelismo implícito en ellos. Según parece, muchos programas de inteligencia artificial se encuadran en este grupo.

Las transformaciones automáticas presentan un inconveniente de índole más fundamental: pueden descubrir el paralelismo latente en algoritmos ya existentes (un algoritmo es el procedimiento por el cual un programa resuelve un problema), pero no pueden inventar algoritmos nuevos, de corte paralelo. Los programadores que se sueltan a pensar en paralelo pueden crear formas totalmente nuevas de resolver problemas. Para plasmar sus inventos en programas operativos precisan lenguajes que permitan expresar el paralelismo, es decir, lenguajes basados en modelos de programación en paralelo.

Por consiguiente, para abordar el paralelismo explícito las soluciones segunda y tercera de mi pregunta inicial se concretan en dos grandes vías. Podemos diferenciarlas por los mecanismos de que se valen para establecer comunicación entre las líneas de actividades paralelas.

La primera escuela imagina un conjunto de actividades paralelas, cada una de las cuales semeja un programa autónomo de tipo convencional. Cada uno de esos "procesos", o subprogramas, consta de estructuras de programa y estructuras de datos; por ejemplo, tablas (*arrays*) y listas, exactamente lo mismo que los programas convencionales. Las estructuras de programa son activas: ejecutan instrucciones para efectuar transformaciones en las estructuras de datos, que son pasivas. (El algoritmo queda plasmado conjuntamente en las estructuras de programa y las estructuras de datos.)

Normalmente, los procesos que colaboran en la resolución de un problema necesitarán compartir datos. Ahora bien, las estructuras de datos se encuentran herméticamente encerradas en cada proceso, sin que unas puedan tomar directamente los datos que necesitan de otras. Por consiguiente, intercambiarán mensajes. Cuando un subprograma tiene datos que otros precisan, genera un mensaje y lo pone en manos de un programa "de correos", que se encarga de encaminar el mensaje hacia su destinatario y entregárselo. Semejante proceder acentúa la complejidad del conjunto del programa, pues obliga a que cada proceso conozca cómo generar mensajes y adónde enviarlos. Pero resuelve el espinoso problema de la coordinación de las comunicaciones entre subprogramas.

Imaginemos que los subprogramas pudieran consultar directamente las estructuras de datos de los demás. Podrían entonces hacer libre uso de datos incompletos o incoherentes, datos todavía en preparación y que no se encuentran listos para que los empleen subprogramas exteriores. Tales problemas de comunicación quedan resueltos al exigir que todas las comunicaciones se efectúen a través de mensajes, pues éstos sólo se generan cuando se dispone de datos completos y provistos de coherencia interna.

La complejidad de casi todos los esquemas de transmisión de mensajes refleja que se adecúan, sobre todo, a los problemas susceptibles de descomposición en unidades bastante grandes e independientes, cuyas necesidades de comunicación puedan reducirse a un intercambio sencillo y habitual de mensajes. Supongamos, por ejemplo, que se trate de simular los movimientos de *n* cuerpos, cada uno de los cuales ejerce una acción gravitatoria sobre lo restantes. Aunque la tarea tenga enunciado sencillo, requiere una enorme computación; en los ordenadores tradicionales, la simulación de sistemas grandes progresa muy despacio. Un buen procedimiento para acelerar la simulación consiste en calcular paralelamente las fuerzas que experimentan la totalidad de los *n* cuerpos, actualizar su posición y repetir el cómputo.

En una técnica descrita para abordar ese problema por Charles L. Seitz, del Instituto de Tecnología de California, se crean *n* subprogramas, uno por objeto. En cada ciclo de cómputo, cada proceso hace circular entre la mitad de los restantes un mensaje que contiene la posición del objeto, al tiempo que recibe mensajes similares de la otra mitad. Imaginemos que un mensaje emitido por P*j* (el proceso encargado del objeto *j*) se detenga en primer lugar en P*k* (el que se ocupa de *k*). P*k* calcula entonces las fuerzas con que en ese momento se atraen *j* y *k*, añade esa información al mensaje y pasa éste a P*l*. Cuando el mensaje retorna por fin a P*j*, porta un informe sobre las fuerzas existentes entre *j* y la mitad de los objetos restantes del sistema. Mientras tanto, P*j* ha sido visitado por mensajes similares procedentes de los restantes objetos del sistema, que le aportan la información que necesita el subprograma para calcular las fuerzas que afectan a *j*. P*j* puede entonces calcular el efecto neto que sobre la posición de *j* ejercen todas las fuerzas.

La pesada tarea impuesta a cada subprograma comporta que, igual que casi todos los problemas fundados en la transmisión de mensajes, semejante técnica de resolución del problema de los *n* cuerpos se adapta especialmente bien a máquinas de "grano" bastante grueso, es decir, máquinas dotadas de procesadores potentes, pero en número muy inferior al de las computadoras de "grano fino". La programación sistólica, una variante radical de la técnica de transmisión de mensajes, es excepción a la regla de grano grueso. Los programas sistólicos prescinden completamente del proceso de entrega de mensajes; funcionan en máquinas cuya red de interconexiones internas ha sido diseñada para el patrón de comunica-

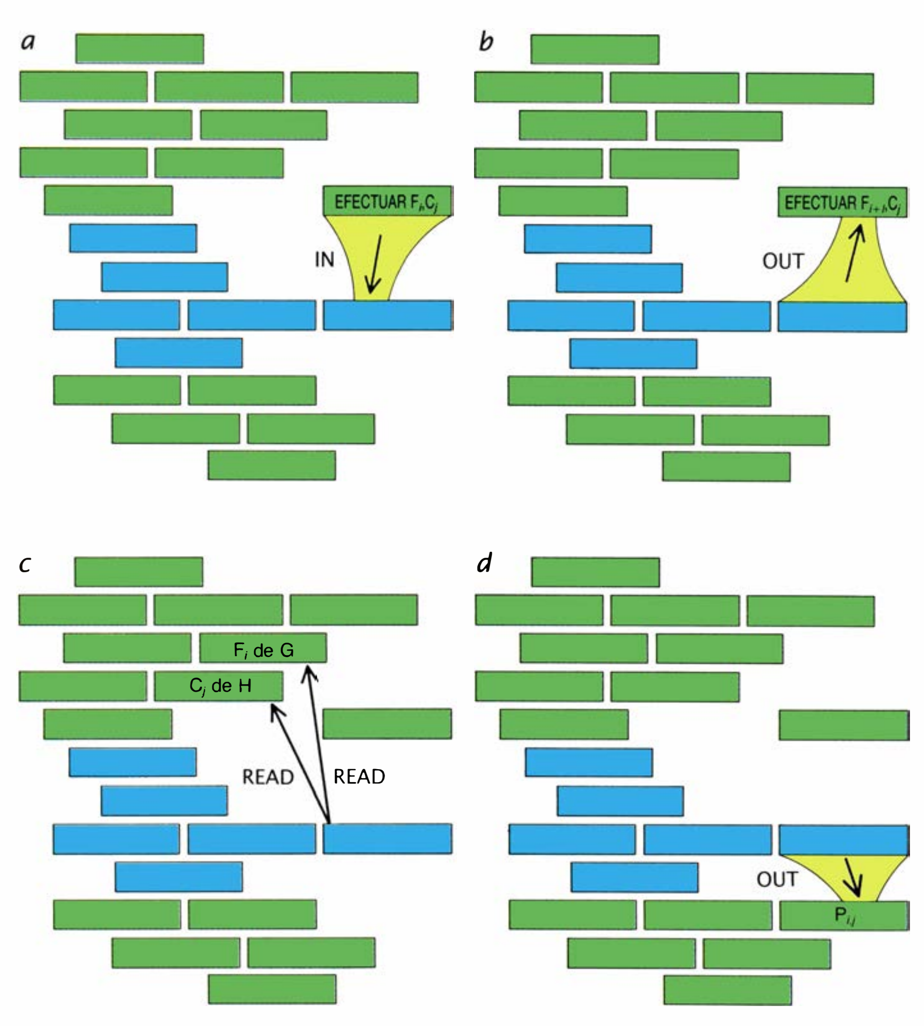
ciones específicamente exigidas por el algoritmo. La organización geométrica de la máquina se corresponde con la organización geométrica del programa. A través de esa red fija de líneas de conexión se van bombeando los datos rítmicamente ("sistólicamente"). Al ser fijos los patrones de comunicación, las computadoras sistólicas son de ordinario máquinas de grano fino, adecuadas a la resolución de problemas compuestos por muchas partes simples. Por ser fija la red de conexiones internas, su campo de aplicación se limita, de ordinario, a objetivos especiales.

En los programas fundados en transmisión de mensajes, cada estructura de datos está herméticamente cerrada en el interior de algún subprograma. En el otro conjunto de métodos de programación paralela, grande y variado, muchos procesos pueden compartir el acceso a una misma estructura. En un ejemplo de ese tipo, una cadena de procesos idénticos, que reciben órdenes de un único proceso "jefe", pueden modificar simultáneamente todos los elementos de una gran estructura. En otro de esos métodos se establecen reglas mediante las cuales muchos procesos acceden, por propia iniciativa, a datos comunes.

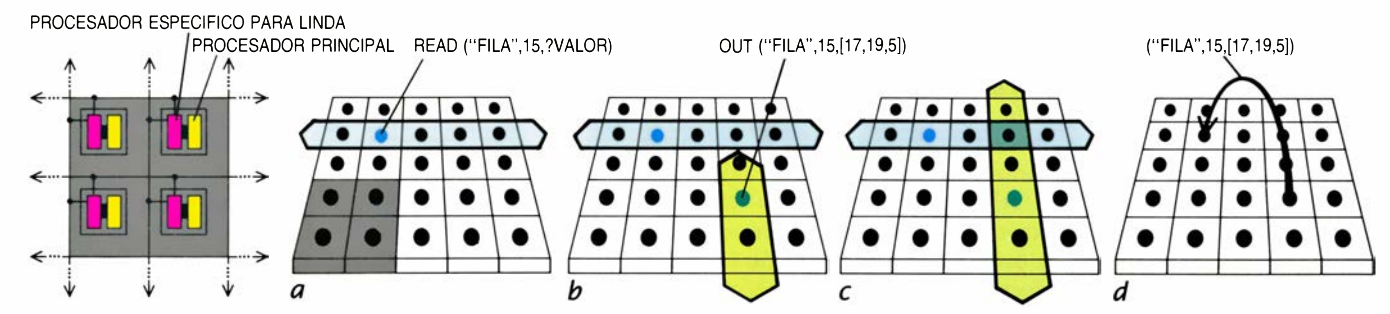
El primero de esos métodos se ejecuta sobre un soporte físico específico; toma cuerpo en varios lenguajes de programación diseñados para una máquina de computación paralela denominada "máquina de conexión" [véase "La máquina de conexión", por W. Daniel Hillis; Investigación y Ciencia, agosto de 1987]. La máquina de conexión es un ordenador en paralelo de grano fino, que puede describirse como una "memoria activa". En el mayor de sus modelos actuales, la memoria de la máquina se divide en unos 65.000 segmentos, controlado cada uno por un procesador sencillo propio. Los programas de la máquina de conexión remedan esa estructura; se fundan en estructuras de datos distribuidas sobre la memoria activa, a razón de un elemento de datos por procesador. Tales estructuras de datos se transforman a sí mismas conforme los procesadores efectúan sincrónicamente, paso a paso, la misma instrucción. Así, por ejemplo, un programa multiplicaría por 3 un conjunto de 10.000 valores por el sencillo expediente de almacenar cada valor en un elemento de memoria activa y emitir la instrucción: "Que cada uno multiplique por 3 su contenido".

El número de procedimientos computacionales que funcionan bien en ese entorno es altísimo; vale de ejemplo la ordenación de una serie de números, una de las operaciones básicas de toda la informática. Uno de los programas de ordenación diseñados para la máquina de conexión se basa en un algoritmo llamado "clasificación rádix". En la versión correspondiente a la máquina de conexión (que se ha deducido de una versión secuencial similar), los números a ordenar de menor a mayor se guardan a razón de uno por procesador, esto es, uno por cada elemento de memoria activa. Todos los procesos revisan simultáneamente el dígito binario menos significativo de la ristra de bits que representan al número en forma binaria. Los números que tienen un 0 en esa posición avanzan un paso en la serie, los que tengan 1 van atrás. (En la máquina de conexión, los números pueden saltar directamente desde cualquier elemento de la memoria activa hasta cualquier otro.) Las posiciones relativas de los números que tengan el mismo bit se conservan durante cada ciclo de la reordenación. El procedimiento se repite para el siguiente dígito binario y, tras un número de iteraciones igual al número de bits del mayor de los números a ordenar, se obtiene una serie ordenada, con el número más pequeño en primer lugar. En tanto no se alcance el límite impuesto por los 65.000 procesadores de que dispone la máquina, el procedimiento ocupa el mismo tiempo, sea cual fuere la cuantía de números a clasificar.

La máquina de conexión permite que le ocurran muchas cosas, al mismo tiempo, a una misma estructura de datos; y coordina toda la actividad sirviéndose de un único subprograma jefe y de una hilera de fieles obreros que, a modo de coristas de ballet, trabajan



6. UN POSIBLE PROGRAMA para multiplicar dos matrices por el método del espacio de tuplas; divide los datos (los contenidos de las matrices a multiplicar, *G* y *H*) en objetos llamados tuplas pasivas (*verdes*) y reparte el trabajo de calcular los productos escalares entre las tuplas activas (*azules*). Para informarse sobre cuál ha de ser el producto escalar que debe calcular seguidamente, una tupla activa consume una tupla pasiva que especifica la fila y la columna pertinentes (*a*). La tupla activa actualiza los datos, dejándolos preparados para la primera tupla activa que busque trabajo, y regenera la tupla pasiva (*b*). Seguidamente busca y lee las tuplas que contienen F_i y C_j, que son la fila y columna a multiplicar (*c*). Tras calcular el producto escalar P_{ij}, la tupla activa genera una nueva tupla pasiva que contiene el resultado (*d*). (Otras tuplas activas se ocupan del cálculo de otros productos escalares.) IN, OUT y READ denotan ciertas operaciones del lenguaje Linda.



7. CREACION DE UN ESPACIO homogéneo de tuplas, en donde cualquier tupla activa pueda tener acceso a cualquier tupla pasiva. Resulta tarea difícil en los ordenadores de funcionamiento paralelo cuya memoria se reparte entre los procesadores. En la máquina Linda, un ordenador paralelo que se está construyendo en los Laboratorios AT&T Bell y que se ha diseñado para funcionar con el lenguaje Linda de manejo del espacio de tuplas, la memoria se encuentra distribuida entre muchos nodos, interconectados en malla cuadriculada (*a la izquierda*). Para leer una tupla pasiva que contenga, por ejemplo, la fila de valores identificada como fila 15, una tupla activa lanza a los restantes nodos de la misma fila una petición READ, acompañada de una descripción parcial de la tupla pasiva (*a*). Puede que otro cómputo se encuentre aún generando la tupla; cuando aparece en algún nodo, éste proyecta su propia señal hacia la parte inferior de su columna (*b*). El nodo donde se intersecan la petición de lectura y la señal expedida por la tupla pasiva (*c*) se encarga de encaminar los datos (*d*).

sincrónicamente en la ejecución de las órdenes recibidas. Otras estructuras de compartición de datos resuelven el problema consintiendo que alguno de los procesos que merodean por la máquina se haga temporalmente con el monopolio de la porción de datos que está manejando e impida que otros procesos actúen al mismo tiempo que él sobre esos mismos datos. A diferencia del estilo pulcramente reglamentado de los programas de la máquina de conexión, esta segunda técnica se adapta bien a problemas con ingredientes heterogénos y con organización más laxa. Además, no se funda en una máquina específica, sino en una noción llamada "espacio de tuplas", que es independiente del equipo material utilizado. El método cobra cuerpo en un lenguaje de programación llamado Linda, desarrollado por Nicholas Carriero, el autor y otros integrantes de nuestro grupo de trabajo de la Universidad de Yale.

Representémonos el espacio de tuplas a modo de un bolsa informe llena de objetos que interactúan entre sí, llamados tuplas, no siendo éstas más que conjuntos ordenados de elementos. Las hay de dos tipos. Las tuplas activas desempeñan el papel de los procesos en el modelo de transmisión de mensajes; son las encargadas de llevar a efecto la computación. Se comunican generando, leyendo y consumiendo tuplas pasivas. Pero a diferencia de los mensajes que reciben y envían los programas de transmisión, las tuplas pasivas forman parte integral de los programas fundados en la noción de espacio de tuplas, pues constituyen las estructuras de datos compartidos.

Durante la ejecución de un programa, las tuplas pasivas se crean y destruyen, cual avisos y notas clavadas con chinchetas en un tablón de anuncios que, una vez caducadas, se retiran. Las tuplas activas que necesitan datos pueden, ora leer tuplas pasivas, ora consumirlas; conforme las tuplas activas van obteniendo resultados, generan nuevas tuplas pasivas (clavan nuevos anuncios en el tablón). Una vez concluido el cómputo que debía realizar, la tupla activa se convierte en tupla pasiva que porta en sí el resultado final de su cómputo. Así, un programa redactado en Linda constituye un enjambre de tuplas activas computando simultáneamente, rodeadas por una niebla de tuplas pasivas, con la que, en última instancia, acaban fundiéndose.

Un programa concebido en Linda para la multiplicación de matrices podría almacenar las filas de una matriz y las columnas de la otra en tuplas pasivas, a razón de una fila o columna por tupla. Una y otra vez, las tuplas activas leen una fila, leen un columna, obtienen su producto escalar y vierten el resultado en una nueva tupla pasiva. Cuantas más tuplas activas haya, tanto más rápidamente se ejecutará el programa (hasta un límite impuesto por la capacidad de la maquinaria de cómputo y por el tamaño del problema). Una solución diferente podría sacar partido de la propiedad de las tuplas activas de pasar a tuplas pasivas cuando concluyen su cómputo: el programa representaría la matriz producto mediante una serie de tuplas activas, responsables cada una de ellas del cálculo de un producto escalar. Las tuplas activas obtienen simultáneamente un producto escalar cada una y se convierten después en el resultado.

Dado que el espacio de tuplas facilita la descomposición de las tareas poco menos que al vuelo, y permite igualmente que el programa cambie de forma durante su ejecución, constituye una prometedora base para la representación de sistemas complejos y sometidos a rápida evolución en un microcosmos. Pensemos, a título estrictamente ilustrativo, en un programa paralelo que modelizase "en tiempo real" (es decir, coincidiendo con el acontecer de los sucesos) la situación actual del tráfico aéreo de los Estados Unidos. Cada vez que despega un avión, se crea una tupla pasiva destinada a mantener información constantemente actualizada que exprese la mejor de las estimaciones posibles del sistema acerca de la situación del aparato. Los datos sobre estado de los aeropuertos, situación meteorológica, etcétera, se guardarían en otras tuplas pasivas.

Las tuplas activas del programa, que operan en paralelo, evalúan esa cambiante imagen y efectúan recomendaciones sobre el control del tráfico. Resulta crucial la capacidad del sistema para concentrar su atención más intensamente en unos lugares que en otros: cuando un sector acuse una actividad particularmente intensa, pueden apiñarse en él muchas tuplas activas, que ayudarán a despejar la situación. Cuando el tráfico se aligera, la tuplas activas pueden dirigir su atención hacia actividades de carácter menos urgente; por ejemplo, la clasificación y recuento de tuplas pasivas que hayan "aterrizado", o la confección de estadísticas. Todavía no se han redactado en Linda programas tan ambiciosos, pero sí se están desarrollando algunos más sencillos de esa misma naturaleza.

¿Cómo alojar una máquina de programa tan fluida en los circuitos, rígidamente interconectados, de un ordenador real? En aquellas máquinas de computación en paralelo donde todos los procesadores están conectados a una memoria común, las tuplas se asientan y guardan en esa memoria compartida. Conforme va siendo necesario para mantenerlos ocupados, a los procesadores se les asignan tuplas activas; cada tupla activa tiene acceso a todas las pasivas. La creación de un espacio de tuplas en máquinas de "me-

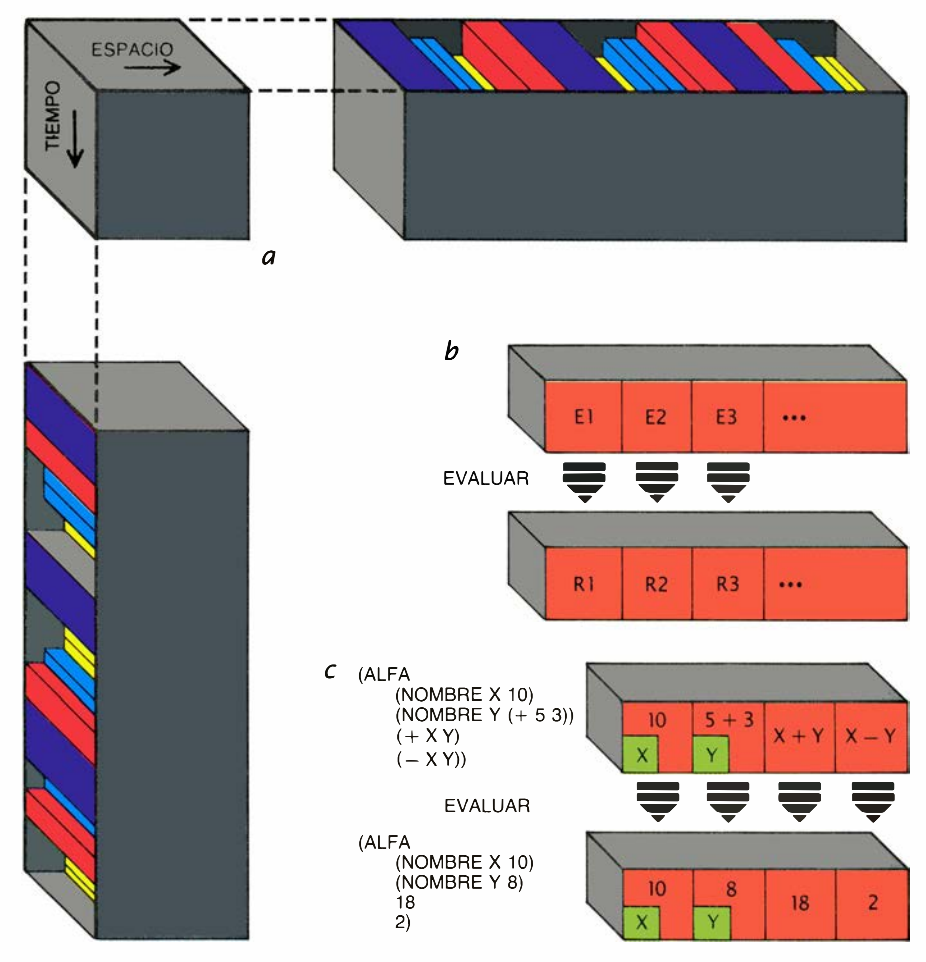
moria distribuida", esto es, repartida entre todos los procesadores sin que ninguna porción de ella la compartan todos, plantea un problema más arduo. Existen, no obstante, sistemas de espacios de tuplas en varias máquinas de ésas, y el grupo de Sudhir R. Ahuja está construyendo en los Laboratorios AT&T Bell una nueva máquina de memoria distribuida proyectada específicamente para utilizarla con Linda.

Todos los modelos de programación que se encuentran en liza poseen alguna ventaja. Según parece, los compiladores paralelizantes sólo funcionan óptimamente sobre una gama bastante restringida de programas; ahora bien, dado que tal clase resulta de suma trascendencia para la computación numérica, los compiladores paralelizantes no cederán importancia. El sistema de transmisión de mensajes es un buen modelo cuando los problemas admiten una descomposición natural en redes lógicas integradas por piezas bien diferenciadas y con alto grado de independencia mutua; sus críticos aducen, empero, que impone un estilo de programación ampuloso y antinatural. Linda comparte (y puede mejorar en este aspecto) la tolerancia del sistema de transmisión de mensajes para habérselas con estructuras de programa diversas y no sincronizadas, y añade por su parte la naturalidad de la compartición de las estructuras de datos. Mas la experiencia disponible con Linda no es todavía suficiente para dar por demostrada la potencia del espacio de tuplas en problemas generales. Si bien el modelo de la máquina de conexión requiere sincronismo y cierta homogeneidad en los problemas, son pocos los modelos de programación capaces de aproximarse siquiera a su capacidad para sacar el máximo partido del uso masivo del paralelismo de grano fino.

Mientras prosigue la investigación sobre modelos operativos de programación, y a pesar de los arreglos y tanteos realizados con los ya existentes, la búsqueda del paralelismo está llevando a los más distinguidos especialistas en ciencias de cómputo a plantearse cuestiones fundamentales sobre la naturaleza de los programas. Las formas de paralelismo que hemos examinado hasta ahora ocupan en realidad una posición periférica respecto del meollo de la computación. Se trata más de un recurso al que echar mano cuando un programa grande y lenturrón necesita que lo espabilen, que de un verdadero instrumento conceptual. Ahora bien, hay otro punto de vista. Si nos encaramamos a ese lugar algo apartado y contemplamos la computación desde lo alto y con la perspectiva adecuada, descubrimos por doquier que en el núcleo de la programación hay cierta clase de paralelismo.

En cierto sentido, resulta incluso imposible concebir un proceso de cómputo que no sea fundamentalmente paralelo. A modo de analogía, pensemos en la receta para hacer una tarta de chocolate. Es corriente que los programas de ordenador se comparen con recetas culinarias: unos y otras se proponen transformar los datos o ingredientes de entrada en alguna clase de producto de salida. De ordinario se piensa que las recetas son fuertemente secuenciales. La receta para hacer una tarta de chocolate podría condensarse así: (1) preparar la masa; (2) añadir las virutas de chocolate; (3) volcar la preparación en un molde; (4) cocerla al horno, y (5) decorar la tarta con merengue. Pero la receta de la tarta es implícitamente (aunque de forma pasiva) una tarea a realizar en paralelo. Para llevarla a la práctica han de existir simultáneamente una serie de objetos: un cuenco para preparar las mezclas, una batidora, un molde, un horno, un cocinero. Pues la existencia de un cuenco, y después de un horno y después de un cocinero, en secuencia de intervalos temporales disjuntos, daría resultados tan insatisfactorios como si el amasado, la cocción y el decorado del pastel se realizaran simultáneamente.

Prácticamente cualquier programa es en esencia paralelo en ese mismo sentido. El programa consta de una colección de piezas; por ejemplo, definición de funciones, de estructuras de datos, de enunciados ejecutables..., las cuales han de existir todas paralelamente para que el programa funcione. ¿Adónde nos lleva semejante observación? En un sentido importante, los cómputos ofrecen desde esa perspectiva el mismo aspecto que desde la tradicional. En otras palabras, parece existir una si-



8. LENGUAJE DE PROGRAMACION SIMETRICO, capaz de reconocer la simetría básica de la estructura de los cómputos (*a*). La ejecución de un programa requiere cierto tiempo y ocupa su "espacio" en el interior del ordenador. Los programas convencionales tratan los cómputos mediante una secuencia temporal de pasos, encajándose los pasos más pequeños en el seno de los grandes. Al examinar el programa bajo una perspectiva espacial, el aspecto del cómputo es muy similar: una colección de objetos coexistentes (módulos, estructuras de datos o definiciones de procedimientos). Un programa de lenguaje simétrico (*b*) especifica la computación mediante un mapa de objetos, que puede tener la sencillez de un valor, la definición de una variable o la complejidad de un gran subprograma. Al ejecutarse el programa, los elementos (*E*) se evalúan en paralelo, para obtener así un nuevo mapa de resultados (*R*). Un programa sencillo en el lenguaje Symmetric Lisp (*c*) especifica una secuencia de objetos, que van transformándose según se ejecuta el programa.

metría fundamental en la estructura de los cómputos.

Tal simetría equivale a lo siguiente. Desde el punto de vista tradicional, casi todos los cómputos constan de una serie de pasos, cada uno de los cuales está definido en función de otra serie de pasos, y así sucesivamente. En la escala máxima, cada paso podría comprender una subrutina compleja; éstas ejecutarían subrutinas más sencillas, encargadas de la evalución de secuencias de enunciados, que se resolverían finalmente en operaciones primitivas de máquina. Desde la perspectiva secuencial, los cómputos resultan, hasta cierto punto, semejantes a sí mismos; la textura de la computación es muy similar cualquiera que sea la escala a que se observe. Los rasgos de gran tamaño se resuelven en rasgos de mediano tamaño, que dan cabida a rasgos menores, y así sucesivamente.

Desde el punto de vista simétrico, esto es, en su aspecto "espacial", los programas están construidos de modo muy similar. En el nivel más grueso suele haber colecciones de módulos, es decir, archivos de código, que probablemente contengan las definiciones de subrutinas y estructuras de datos, que pueden a su vez contener declaraciones de grano más fino, y así sucesivamente, hasta descender a valores primitivos y a variables y expresiones ejecutables (y en última instancia, a palabras de la memoria). Los lenguajes de programación convencionales sacan partido de tal sibisimilitud temporal y permiten al programador tratar como equivalentes pasos de cualquier tamaño, con lo cual pueden encadenarse de igual modo dos operaciones simples o dos operaciones complejas, y un mismo tipo de bucle puede, por ejemplo, ejecutar por iteración un millón de operaciones o una docena. Sin embargo, la sibisimilitud paralela, o "espacial", queda oculta, y los programadores se ven obligados a definir de modo distinto objetos de diferentes escalas.

Existe una nueva clase de lenguajes de programación, que admite la estructuración sibisimilar tanto a lo largo del eje espacial como del temporal. Estos lenguajes "simétricos", que nuestro grupo está desarrollando en colaboración con Suresh Jagannathan, del Instituto de Tecnología de Massachusetts, Thomas London, de los Laboratorios Bell, y otros colegas, permiten a los programadores descomponer los cómputos no sólo en sucesiones temporales de pasos cada vez más simples, sino también en secuencias paralelas, o "es-

paciales", compuestas por objetos cada vez más sencillos. En los programas redactados en lenguajes simétricos, los objetos paralelos, que en los programas convencionales quedaban ocultos, se encuentran ahora a la vista en el "espacio" interior del ordenador. Se hallan dispuestos en configuración espacial bien definida (sabemos dónde se aloja cada objeto, exactamente lo mismo que sabemos en qué momento se da cada uno de los pasos de una serie temporal); además, dentro de cada objeto se pueden encajar objetos adicionales, dispuestos, lo mismo que antes, en configuración bien definida. Los objetos pueden aludir unos a otros tanto por su posición como por su nombre. Así, si denomináramos x un objeto que contuviera el número 3, el contenido de otro objeto podría darse como $x + 5$, o tal vez, como "el contenido del decimoséptimo objeto contando desde la izquierda" + 5. Por tanto, el texto de un programa redactado en lenguaje simétrico describe el contenido de cada objeto y su posición sobre un mapa.

Cuando se ejecuta tal programa, todas las regiones se evalúan simultáneamente y el resultado es un nuevo mapa que refleja el estado final de cada objeto, correspondiéndose biunívocamente con el mapa inicial de expresiones no computadas. Durante su ejecución, el programa simétrico nos recuerda el proceso de revelado de un negativo fotográfico.

Lo que aprendamos sobre la construcción de máquinas-programa en paralelo puede, en última instancia, cambiar nuestra forma de comprender tanto la programación como los límites de lo computable. Los programas en paralelo son (en principio) inmensamente más potentes que los programas secuenciales, a punto de pasar de moda. ¿Adónde nos conduce todo ello? El paralelismo transformará muchos campos de actividad, desde los gráficos y los programas de supervisión y control "en tiempo real" hasta la inteligencia artificial y mucho más. Y, lo que quizá sea más importante, el paralelismo permitirá modelizar y analizar fenómenos naturales con un nivel de perfección muy superior al actual. Martin H. Schultz opina que el objetivo final del paralelismo es "la eliminación de la ciencia experimental". La afirmación de Schultz no debe interpretarse en toda su literalidad —pues sí necesitaremos una ligera dosis de datos experimentales, para depurar nuestros modelos computarizados—, pero debería constituir motivo de reflexión.



Almacenamiento de datos para computación avanzada

De aquí a cinco años, los dispositivos magnéticos duplicarán la velocidad de escritura y lectura de datos en unos discos que contendrán el quíntuple de la información actual. Los ópticos ofrecerán mayor densidad de almacenamiento

Mark H. Kryder

Sabido es que el valle de Santa Clara, en California, constituye el centro de la industria electrónica de los Estados Unidos. De ahí su apodo "Silicon Valley" (Valle del Silicio); ese apelativo molesta a los fabricantes de dispositivos magnéticos de almacenamiento de datos. Después de todo, las empresas establecidas en el valle consiguen hoy la mayoría de sus ingresos de los dispositivos magnéticos, y no de los semiconductores. Mejor sería, sugieren los fabricantes, llamarlo "Valle del óxido de hierro", pues de este material están hechos los medios de grabación magnéticos.

Puntillos aparte, la tecnología magnética de almacenamiento de datos, en cuanto pieza importante de los dispositivos de disco flexible y cinta magnética y de los dispositivos de disco duro, más refinados, resulta tan decisiva para los ordenadores actuales como la tecnología de semiconductores. Los poderosos sistemas de computación del mañana sólo serán útiles si acceden a un mayor volumen de datos y en menos tiempo que los sistemas actuales. De aquí que la mayor velocidad de computación ofrecida por los avances en los dispositivos semiconductores y por la interconexión y programación de éstos deba ir acompañada de un aumento en la capacidad y prestaciones de los dispositivos de almacenamiento de datos.

Los dispositivos magnéticos permiten el almacenamiento permanente de una enorme cantidad de datos, pues ofrecen una capacidad de memoria mucho mayor y a un coste por bit de dato almacenado más bajo que los dispositivos semiconductores. Un dispositivo de disco duro típico almacena 800.000 caracteres (algo así como 400 páginas mecanografiadas) en un centímetro cuadrado del medio de grabación; algunos de los mayores dispositivos de multidisco almacenan más de 5000 millones de caracteres. Los datos pueden hallarse en 15 milésimas de segundo y recuperarse a una velocidad de 3 millones de caracteres por segundo. Además, puesto que algunos discos y todas las cintas pueden sustituirse fácilmente sin destruir los datos que almacenan, constituyen un archivo de información casi ilimitado.

No solamente son impresionantes las capacidades de almacenamiento y búsqueda de esta tecnología magnética, sino que también algunos de sus equipos representan una ingeniería consumada. El dispositivo de disco duro consta de una placa extremadamente lisa recubierta con un medio magnético de grabación; gira a 3600 revoluciones por minuto. Un cabezal de grabación microscópica —volando literalmente sobre una almohada de aire de 10 millonésimas de pulgada de espesor— se desliza sobre la superficie de la placa a una velocidad relativa de unos 45 metros por segundo; procede circularmente siguiendo una pista de media milésima de pulgada de anchura. (Una pulgada equivale a 25,4 milímetros.) El cabezal está sometido a muchas veces la aceleración de la gravedad terrestre en su movimiento veloz de pista a pista; aún así, se las arregla para encontrar y alcanzar su destino sin errores sobre la pista prevista. Si el cabezal "aterrizase" sobre el medio a semejante velocidad, el disco y toda la información recogida quedarían destruidos. Pero no ocurre tal, y estos dispositivos alcanzan una vida útil de más de 10 años.

A pesar del magnífico diseño de los actuales dispositivos magnéticos de almacenamiento de datos, se prevé, para el futuro, mejores prestaciones. Se espera que nuevos materiales magnéticos y cabezales de grabacion más avanzados conduzcan, de aquí a cinco años, a mejoras que quintupliquen la densidad de almacenamiento conocida por los productos actuales. Muchos de los dispositivos continuarán "escribiendo" (almacenaje) y "leyendo" (búsqueda) datos por medio de cabezales electromagnéticos de grabación, como los de los magnetófonos de grabación sonora; pero los dispositivos que utilicen láseres para leer, escribir y borrar la información ocuparán un segmento notable del mercado de almacenamiento de datos. Tales grabaciones magneto-ópticas alcanzarán densidades de almacenamiento de datos más altas y sustituirán con ventaja la grabación magnética tradicional en múltiples aplicaciones.

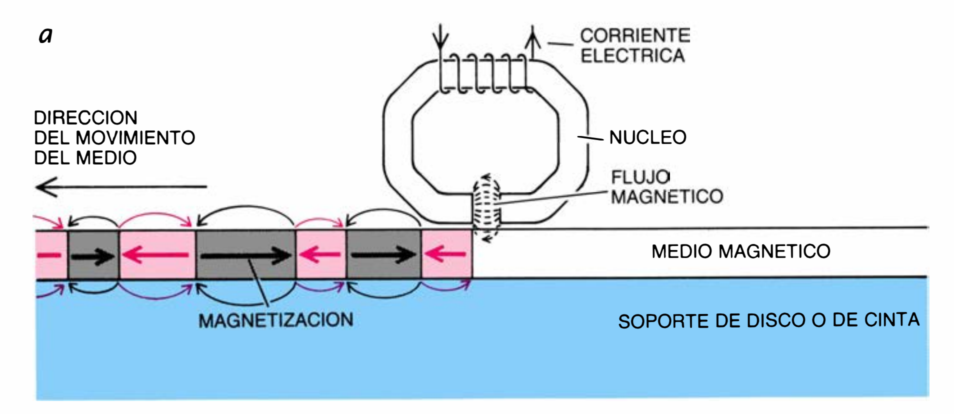
Dejando de lado la distinción entre cintas o discos, el almacenamiento magnético de datos binarios consiste, en definitiva, en imprimir un patrón de magnetización sobre un medio. Ello corre a cargo de un cabezal de grabación inductiva, que consta de un filamento conductor bobinado varias veces alrededor de un núcleo anular que puede magnetizarse fácilmente. La corriente eléctrica del conductor induce un flujo magnético en el núcleo; invirtiendo el sentido de la corriente se invierte la dirección del flujo magnético. Puesto que las líneas de flujo magnético se propagan a medida que van sal-

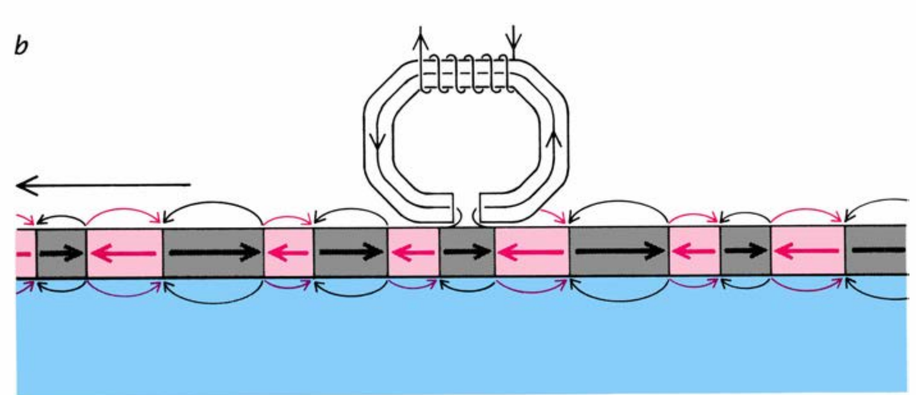
1. INTEGRACION DE DOS TIPOS DE CABEZALES DE LAMINA DELGADA en un dispositivo microscópico fabricado por la empresa PCI del holding Central Data Corporation. La estructura bobinada marrón, similar a la bobina de hilo conductor de un cabezal inductivo tradicional, conduce una corriente que establece un flujo magnético a través del entrehierro ("gap") para escribir datos sobre un medio. Un cabezal magnetorresistivo inferior, que detecta campos magnéticos como cambios de resistencia, lee los datos. Tal cabezal es sensible a los campos por encima mismo de una pista de datos. Al poder construirse más estrecho que la pista, el elemento de lectura resulta ser también menos susceptible a las interferencias.

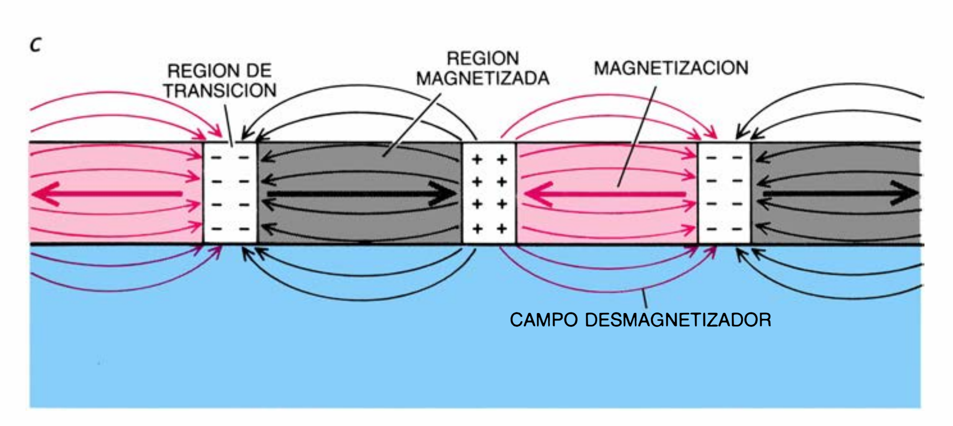
vando intervalos del núcleo anular (también llamados "entrehierros" o "gaps"), pueden servir para magnetizar un medio situado en la vecindad del entrehierro. La dirección de la magnetización depende de la dirección de las líneas de flujo magnético que "puentean" el entrehierro.

Así pues, un patrón de magnetización se irá imprimiendo en el medio, conforme éste se mueve bajo el entrehierro, con la mera inversión repetitiva del flujo de corriente del cabezal de grabación. Ahora bien, los datos generados por computador se almacenan en cadena de dígitos binarios; en consecuencia, el patrón de inversiones de corriente (y por ende el patrón de inversiones de magnetización) corresponderá a una sucesión de ceros y unos que constituyen los bits de datos.

Para leer los datos codificados en un patrón de magnetizacion del medio, se aprovechan los campos magnéticos establecidos por cada región magnetizada. Estos campos actúan de forma opuesta a la magnetización de la región [*véase la figura 2*], por cuyo motivo se les denomina campos desmagnetizadores. Puesto que las líneas del campo desmagnetizador se extienden algo más allá de la superficie del medio, pueden intersecar el cabezal e inducir un flujo magnético en el núcleo, en tanto en cuanto el cabezal permanezca muy cerca del medio. El proceso de lectura se realiza invirtiendo el proceso de grabación: cuando el cabezal pasa por una región de transición, que se caracteriza porque la magnetización del medio cambia de dirección, el cambio correspondiente en la cantidad de flujo que atraviesa el núcleo se registra como una corriente eléctrica inducida en el conductor que lo rodea. La variación de corriente eléctrica inducida puede detectarse e interpretarse en forma de bits de datos.







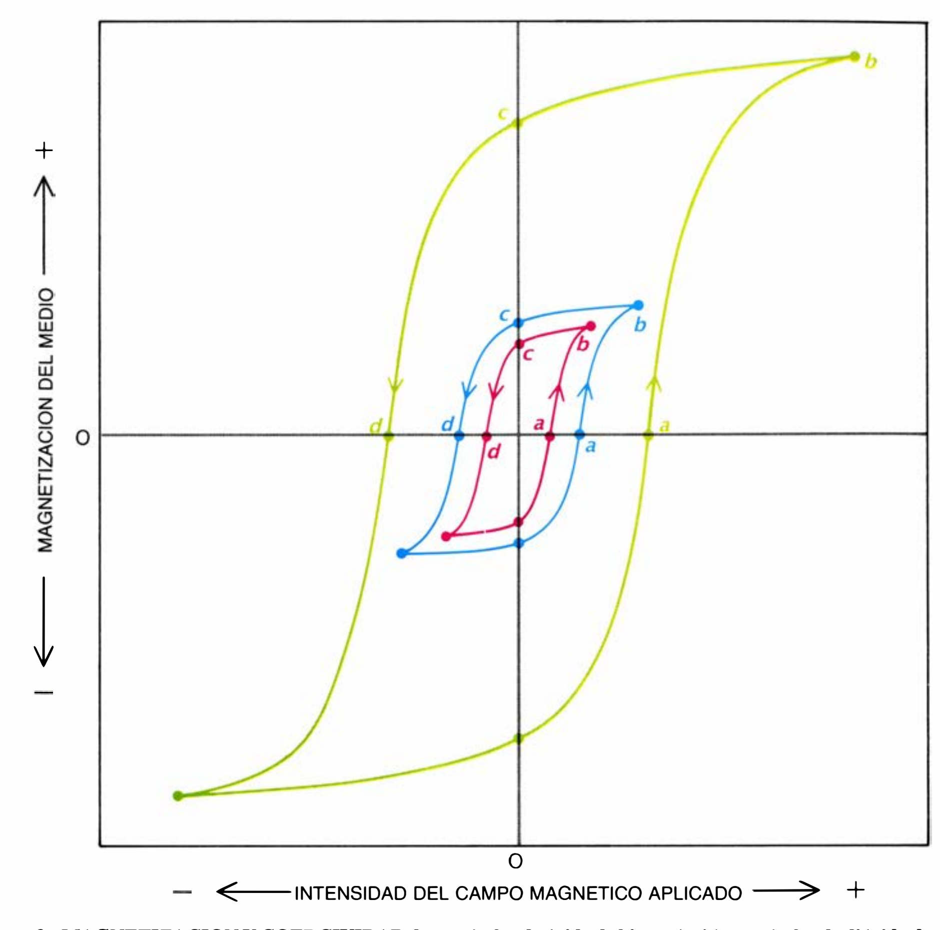
2. ALMACENAMIENTO MAGNETICO DE DATOS mediante la magnetización de zonas de un medio suprayacente a un disco en movimiento o a un soporte de cinta. Esas regiones se magnetizan, en uno de dos sentidos opuestos, por un cabezal electromagnético; consta éste de un núcleo fácilmente magnetizable rodeado de una bobina conductora (*a*). Cuando la corriente fluye por la bobina, se genera en el núcleo un flujo magnético. Conforme el flujo del núcleo salva un intervalo (también llamado entrehierro o "gap"), las líneas de flujo se esparcen y magnetizan el medio subyacente a su paso. Si la corriente se invierte, el flujo y el sentido de magnetización del medio se invierten también. Los datos así almacenados pueden ser "leídos" gracias a que los campos magnéticos generados por las regiones magnetizadas trascienden la superficie del medio. Durante el deslizamiento del medio, el núcleo del cabezal queda expuesto a un flujo magnético cambiante, que induce una corriente en la bobina (*b*). El patrón de regiones magnetizadas y sus campos (*c*) nos manifiesta que los campos comienzan y terminan entre esas zonas, en las llamadas regiones de transición. Por oponerse tales campos a la magnetización del medio, se les denomina habitualmente campos desmagnetizadores.

Si se trata de un disco, los bits de datos se graban normalmente en pistas circulares concéntricas. Hasta cierto punto, el número de bits que caben en una pista (la densidad lineal de bits) puede aumentarse menguando las regiones de transición y, por consiguiente, facilitando la introducción de más regiones magnetizadas en la pista. A pesar de ello, la dimensión mínima de una región de transición tiene un límite. Las regiones magnetizadas de una pista suelen representarse con la imagen de un tren de barras, colocadas una a continuación de otra. La región de transición separa cada uno de los polos de las regiones magnetizadas (norte y sur) de los polos similares de las regiones vecinas. Menguar la región de transición implicaría colocar muy próximos los polos magnéticos iguales, disposición desechada puesto que polos idénticos se repelen. Además, si las regiones magnetizadas se hacen demasiado pequeñas en relación al tamaño de las regiones de transición, los campos desmagnetizadores pueden adquirir intensidad suficiente para invertir la magnetización de una región. El dilema se desvanece asegurando que el medio tenga alta coercividad: una medida de la magnitud del campo magnético que debe aplicarse para provocar que la magnetización invierta su dirección en el medio.

Otra forma de aumentar la densidad lineal de bits en una pista consiste en estrechar la capa del medio magnético. Puesto que los campos desmagnetizadores surgen de las regiones de transición, el "adelgazamiento" de las regiones debilita los campos. Si el medio se recorta en demasía, los campos desmagnetizadores que lo trasciendan re-

sultarán muy débiles y no se detectarán, imposibilitando de ese modo la lectura de los datos codificados en el patrón de magnetización del medio.

De nuevo, una elección cuidadosa del material del medio nos aliviará el problema. Si un material puede convertirse en intensamente magnético, no se necesitará que se deposite gran cantidad del mismo sobre el sustrato. Cabe, pues, contar con una capa de medio más delgada mientras se mantiene la misma amplitud de señal de lectura en el cabezal. De los materiales intensamente magnéticos se dice que presentan una alta magnetización. Se infiere, de lo anterior, que ambas propiedades —una alta magnetización y una alta coercividad— sean las deseables en un medio magnético, si no quiere perderse la señal de lectura cuando crece la densidad lineal de bits.



3. MAGNETIZACION Y COERCIVIDAD de partículas de óxido de hierro (*rojo*), partículas de dióxido de cromo (*azul*) y aleaciones de lámina delgada (*verde*), indicadas, respectivamente, por la altura y anchura de cada bucle de histéresis: el camino trazado por la magnetización de cada medio en función del campo magnético aplicado. Los bucles de histéresis demuestran que, conforme aumenta la intensidad de campo, aumenta también la magnetización desde cero (*a*) hasta un valor máximo (*b*). Si después se elimina el campo, la magnetización no vuelve a cero (*c*); sólo se anulará si se invierte el campo magnético y se eleva hasta el valor de coercividad del material (*d*). La coercividad mide la resistencia del material a invertir su magnetización. Para servir de almacenamiento de datos de alta densidad debe mantener una alta magnetización y poseer una coercividad elevada; en otras palabras, el bucle de histéresis del material será largo y ancho. En razón de este criterio, las aleaciones de lámina delgada constituyen los mejores medios para alta densidad.

Las partículas aciculares de óxido de hierro (en su forma de óxido de hierro gamma) que han sido inmovilizadas en una suerte de matriz cubierta por ambos lados gozan de las dos propiedades deseadas: alta magnetización y coercividad. Por ello, ese material compuesto se ha convertido en el medio de grabación más difundido. Para elevar la coercividad del medio de óxido de hierro todavía más, suelen recubrirse las partículas con cobalto. Otro medio de alta coercividad consiste en partículas de dióxido de cromo en una matriz similar a la anterior. Las partículas de dióxido de cromo tienen una coercividad similar a la de las partículas de óxido de hierro recubiertas con cobalto; presentan una forma más regular que facilita, así se cree, su dispersión en la matriz, de lo que resulta un medio mejor de grabación.

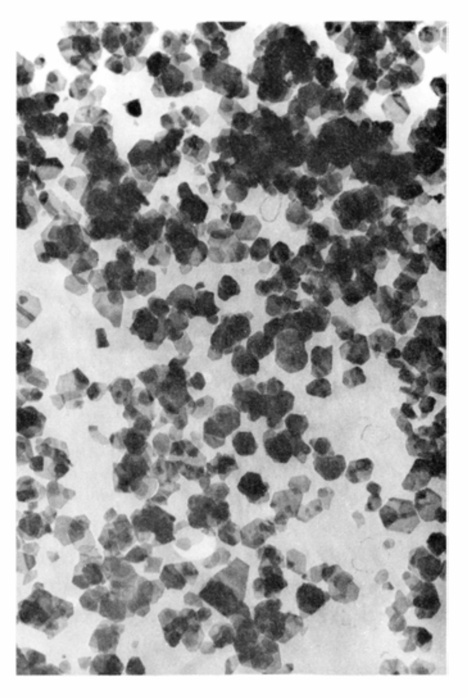
Para grabaciones avanzadas de alta densidad se está ensayando con partículas de metal puro, hierro, por ejemplo. Ofrecen una coercividad y magnetizacion más altas que las partículas de los medios magnéticos actuales. El problema que las partículas de hierro puro plantean es que son tan magnéticas y tan pequeñas que se hace difícil separar unas de otras en detrimento de su deseada dispersión uniforme en el aglutinador. Además, el hierro puro se oxida; las partículas finas de hierro arderían en cuanto se las expusieran al aire; de ahí que requieran un extremado cuidado en la manipulación a que se ven sometidas durante la fase de fabricación. Por otro lado, el recubrimiento de la matriz no solamente debe mantener las partículas sobre el sustrato, sino actuar también de barrera contra el vapor de agua y el oxígeno. Por último, el hierro puro, más blando que el óxido de hierro, se muestra, asimismo, más sensible al desgaste.

Aunque los medios de partículas continuarán empleándose en dispositivos magnéticos de almacenamiento de datos, las láminas finas de aleaciones metálicas, que contienen principalmente níquel y cobalto, acaban de entrar con paso decidido en el mercado de dispositivos magnéticos avanzados de almacenamiento de datos. Los medios de láminas finas no están hechos de partículas, sino que son policristalinos: construidos por pequeñísimos cristales, o granos, de material magnético. Los granos se comportan como si fueran partículas magnéticas mínimas, aunque empaquetadas en una densidad muy superior a como se disponen las partículas en los medios de grabación ordinarios. Pueden conseguirse, pues, medios de láminas finas que tengan coercividades más altas que los constituidos de partículas y ofrecer, a la vez, magnetizaciones más intensas.

La mayoría de los medios de láminas finas se depositan sobre un sustrato en forma de disco a través de un proceso de recubrimiento. Se sumerge el disco en un baño químico que alberga, entre otras sustancias, el material que va a constituir el medio. La composición del baño, la temperatura y su agitación determinarán la composición final y las propiedades del medio depositado.

Para depositar los medios de láminas finas sobre un sustrato se recurre también a otro proceso: pulverización de partículas por bombardeo iónico ("sputtering"). Esa pulverización se realiza en una cámara de baja presión que contiene un gas inerte, argón por ejemplo. En la cámara hay un electrodo cargado negativamente, llamado cátodo, y un electrodo mantenido a potencial de tierra o muy próximo, llamado ánodo. El material a depositar sobre el sustrato se coloca en contacto eléctrico con el cátodo. El sustrato (generalmente un disco) se mantiene casi a potencial cero y se coloca en la ve-





4. PARTICULAS MAGNETICAS embebidas en un aglutinador similar a un pegamento. Constituyen el medio magnético más común para almacenamiento de datos. Los medios particulados habituales contienen la forma gamma del óxido de hierro (*izquierda*). Algunos fabricantes sostienen que, debido a su forma más regular, las partículas de dióxido de cromo (*centro*) se esparcen con mayor uniformidad en el aglutinador. Ambos compuestos muestran una dirección preferida de magnetización a lo largo del eje de la partícula. Las partículas de ferrita de bario (*derecha*) semejan, por su forma, placas hexagonales. Siendo perpendicular a las placas la dirección de magnetización preferida, se ha pensado en el compuesto para su aplicación en medios de grabación perpendicular.

cindad del material cargado. Un filamento caliente ioniza el argón de la cámara de bombardeo; la alta tensión negativa (1000 volts o más) del cátodo y del material unido a él atrae los iones de argón cargados positivamente. Los iones de argón chocan con el material con tal energía que le arranca átomos o racimos de átomos, que caen sobre el sustrato de las inmediaciones.

Hasta el momento, las láminas delgadas obtenidas por deposición química han gozado de mayor difusión que las creadas por pulverización iónica, lo que cabe atribuir, en parte, al coste mayor del segundo proceso. Pero los medios fabricados según el método de pulverización de partículas por bombardeo iónico podrían sustituir a los medios de deposición química, andado el tiempo, ya que resulta más fácil añadir cantidades pequeñas de cromo o renio, elementos que aumentan la coercividad.

Además de una selección cuidadosa de las propiedades magnéticas del medio, los cambios operados en el diseño del cabezal de lectura-escritura ofrecen también la posibilidad de mejorar las prestaciones de los sistemas de almacenamiento de datos. En particular, la instalación del cabezal más cerca del medio —reducción de la llamada altura de vuelo del cabezal— permitiría el registro de señales magnéticas más fuertes, tanto en el medio como en el cabezal. Estas señales más fuertes permitirían, a su vez, que el dispositivo funcionara con mayor rapidez y sin errores. Ahora bien, resulta dificilísimo reducir la altura de vuelo de los dispositivos de disco magnético. El espacio entre el medio y el cabezal en los discos duros se ha restringido tanto que hasta la mínima contaminación por polvo, huellas dactilares o partículas de humo provocan el "aterrizaje" del cabezal destruyendo los datos del disco. (Por esa razón los discos duros se diseñan de suerte tal que no puedan extraerse de sus compartimentos.)

Incluso en las unidades de grabación de cinta y en las de discos flexibles, reputadas generalmente como tecnologías de grabación de "contacto", se deja una separación mínima entre el cabezal de grabación y el medio magnético. Y así, se aplica una lámina de lubricante sobre el medio para evitar el excesivo desgaste del cabezal y el medio. La propia rugosidad de la superficie del medio y su velocidad relativa aseguran que el cabezal "rebote" realmente a lo largo del medio, entrando sólo en contacto ocasional con él. En principio, la separación podría acortarse en los dispositivos de grabación de contacto aplicando una presión mayor entre el cabezal y el medio, pero estas medidas comportarían un mayor desgaste de ambos.

Para aumentar la densidad lineal de bits se puede introducir otro cambio en el diseño del cabezal: estrechar el entrehierro en el núcleo de la cabeza. Esa menor anchura del entrehierro reduce el tamaño de las regiones magnetizadas en el medio, y permite grabar más sobre cada pista. El núcleo de la mayoría de los cabezales de grabación magnética está hecho de ferrita, material que contiene óxidos de hierro y otros metales (manganeso, zinc y níquel). En los últimos cinco años las anchuras del entrehierro de los cabezales de ferrita se han acortado de 2,5 micras a 0,5 micras. Desde hace poco se fabrican cabezales de grabación avanzados sobre sustratos cerámicos con las mismas técnicas empleadas en la producción de dispositivos semiconductores microscópicos. Tales cabezales miniaturizados, hechos de una lámina fina de Permalloy (aleación de níquel y hierro), suelen presentar anchuras de entrehierro de media micra. Los entrehierros de futuros cabezales de lámina fina dividirán por dos ese valor.

Hay cambios en el diseño del cabezal, que, si bien no aumentan la densidad de bit por pista, pueden conseguir una densidad total de almacenamiento de datos mayor. El cambio más obvio sería reducir la anchura del cabezal, pues ésta determina la anchura de la pista. La construcción de pistas más finas permite un mayor número de ellas en el disco, pero también hace más difícil mantener la cabeza sobre la pista correcta y evitar interferencias de las señales magnéticas de las pistas adyacentes. Eso significa que, aceptando

que la limitación de la anchura de la pista sea el cambio más probable para alcanzar densidades de grabación superficiales (por oposición a lineales) más altas, ésas sólo se producirán con el desarrollo de cabezales y mecanismos de seguimiento de pistas más complejos. La tendencia hacia estructuras de cabezal más complejas acelerará el tránsito de los cabezales de ferrita a los cabezales de lámina fina.

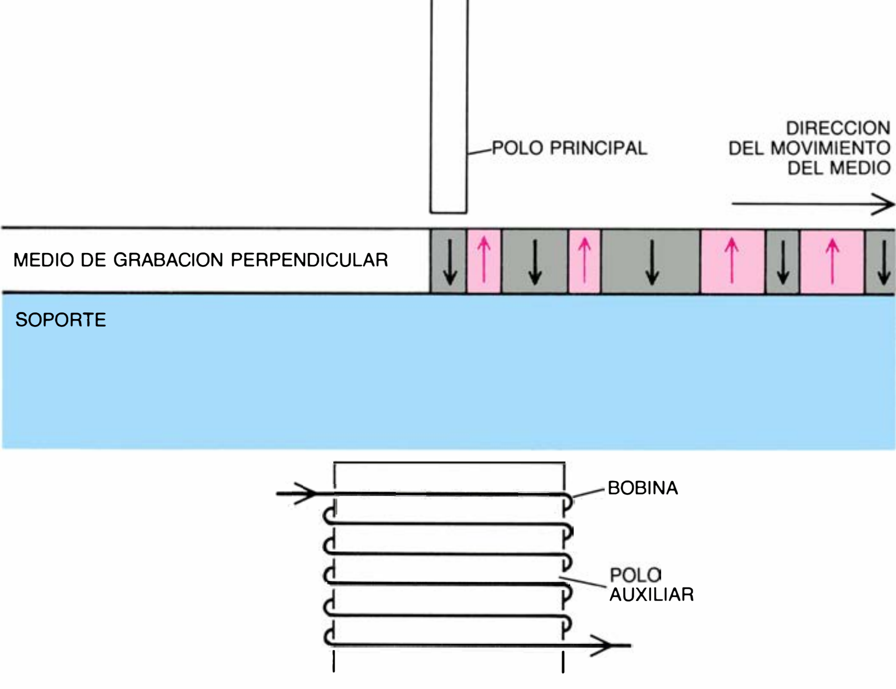
Han dado ya un paso en esa dirección las compañías IBM y Control Data Corporation con la introducción de un dispositivo que combina un cabezal inductivo normal para escritura y un nuevo tipo para lectura: un cabezal magnetorresistivo. El material magnetorresistivo varía su resistividad eléctrica en función de su magnetización. Por consiguiente, los cambios introducidos en la magnetización del medio pueden modificar la resistencia de un cabezal superpuesto que sea del material en cuestión. Si se aplica una corriente constante al cabezal, el cambio en la resistencia puede detectarse de inmediato como un cambio de tensión. Los cabezales magnetorresistivos ofrecen, como mínimo, un orden de magnitud mayor, en sensibilidad al flujo magnético que emana del medio, que el valor conseguido por los cabezales inductivos. Además, su señal de salida depende del flujo total, no de la tasa de cambio de dicho flujo, como ocurre con los cabezales inductivos. Mientras que en un cabezal inductivo la señal eléctrica depende de la velocidad con que el medio cruza delante suyo, la señal de un cabezal magnetorresistivo resulta ser independiente de la velocidad del medio con respecto al cabezal. De ahí las ventajas que ofrecen los cabezales magnetorresistivos, siempre que la velocidad relativa del medio sea baja o variable, que es lo que suele ocurrir en las unidades de cinta.

Los cabezales magnetorresistivos no pueden generar ningún campo magnético externo; se limitan a leer datos. Por consiguiente, el cabezal magnetoresistivo debe ir siempre acoplado con un cabezal inductivo que lea los datos. A pesar de lo cual no se requiere que el cabezal magnetorresistivo sea una estructura separada. En razón de su finura (entre 0,03 y 0,05 micras), los cabezales magnetorresistivos pueden insertarse en el entrehierro de una ferrita e incluso en el de un cabezal grabador inductivo de lámina delgada. La integración de un cabezal de lectura magnetorresistivo con un cabezal de escritura inductivo, además de mejorar la sensibilidad de retrolectura, resuelve los problemas de interferencia entre pistas durante el proceso de lectura, merced a que la anchura del elemento de lectura puede restringirse más que la anchura de la pista.

Estrictamente hablando, lo examinado hasta aquí se aplica a la grabación magnética longitudinal: la lectura y escritura de datos grabados como cambios en la magnetización paralela al plano de un disco o una cinta. Puede conseguirse también que la dirección preferida de magnetización de ciertos medios sea perpendicular al plano de éstos. Los medios habituales de grabación perpendicular son aleaciones de cobalto y cromo depositadas por pulveración iónica ("sputtering"). Estos medios policristalinos tienden a presentar todos sus granos orientados de suerte que la dirección preferida de magnetización neta sea perpendicular al plano de la película. Desafortunadamente, tales películas finas no se distinguen por su oposición al desgaste. Por esto, la ferrita de bario —un medio de partículas— se está también investigando para la grabación perpendicular. Las partículas lenticulares de ferrita de bario pueden mezclarse con partículas de extremada resistencia contra la abrasión y con lubricantes para minimizar el desgaste del medio.

Se han presentado varios diseños de cabezales para grabación perpendicular. De acuerdo con uno de los mismos, el cabezal consta de dos polos separados: un polo auxiliar ancho ubicado en un lado del medio y un polo principal fino colocado cerca del medio, aunque por el otro lado. El polo auxiliar porta un conductor rebobinado que genera un flujo magnético en el polo siempre que lo atraviesa una corriente eléctrica. El campo magnético del polo auxiliar, aunque inadecuado para magnetizar el medio por sí mismo, puede magnetizar el polo principal de la otra cara del medio. El campo magnético combinado de los dos polos adquiere entonces intensidad suficiente para provocar que el medio se magnetice perpendicularmente; por así decirlo, las regiones magnetizadas del medio están apiladas. El patrón codificado de tales regiones perpendicularmente magnetizadas puede leerlo el cabezal de dos polos, igual que un cabezal convencional lee un patrón de regiones magnetizadas longitudinalmente. Los cambios producidos en la dirección de los campos magnéticos del medio inducen un flujo magnético variable en el polo auxiliar. La variación del flujo se detecta entonces como inversiones de corriente en el arrollamiento del polo.

Ocurre que las regiones magnetizadas antagónicamente en los medios de grabación perpendicular no tienen sus extremos uno frente al otro (sí se da eso en los medios de grabación longitudinal); por consiguiente, los campos

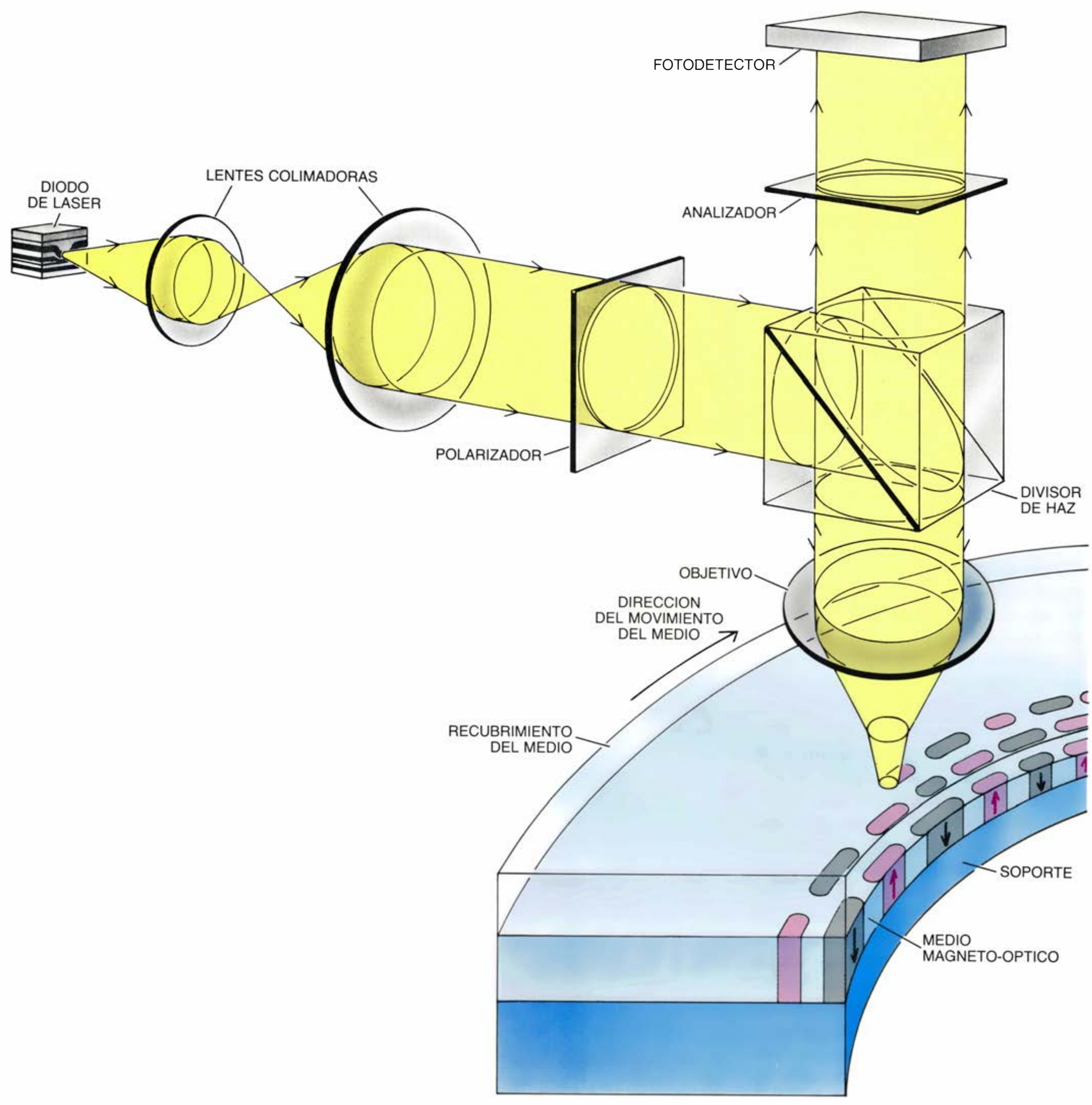


5. GRABACION PERPENDICULAR mediante dos polos colocados sobre lados opuestos del medio magnético. Cuando la corriente fluye por la bobina del polo auxiliar grande, se induce un flujo magnético en el polo. Este campo, aunque insuficiente para magnetizar el medio, sí puede magnetizar el polo principal del otro lado del medio. El medio se magnetiza entonces perpendicularmente a su dirección de desplazamiento por los campos combinados de ambos polos. Puesto que las regiones magnetizadas no quedan colocadas extremo con extremo como ocurre con las grabaciones de datos ordinarias, las regiones de transición entre ellas resultan más delgadas, permitiendo que una pista del medio contenga más regiones magnetizadas.

desmagnetizadores en los medios de grabación perpendicular tienden a ser más débiles conforme aumenta la densidad lineal de bits. Otra ventaja de la grabación perpendicular, nacida de la disposición diferente de las regiones magnetizadas, reside en el medio: podemos aumentar su grosor, y por tanto convertirlo en más resistente a los defectos, sin tener que incrementar el tamaño de las regiones de transición. Y por ser más débiles los campos desmagnetizadores, existe una menor variabilidad en el espaciado y anchura de las regiones magnetizadas, lo que comporta menos errores de lectura y escritura.

A pesar de sus ventajas inherentes, la grabación perpendicular no ha entrado todavía en la fase de comercialización. Ello obedece, en parte, a que, para beneficiarse de las densidades lineales de bits potencialmente más altas, será probablemente necesario colocar el polo principal más cerca del medio, como ocurre ahora para los cabezales de grabación longitudinal. Pero, según se explicó antes, nos hallamos ahí con una barrera casi infranqueable en nuestro deseo de alcanzar densidades lineales de bits más altas en cualquier tipo de grabación magnética.

La comercialización de la grabación perpendicular encuentra su principal obstáculo en la imposibilidad de introducirla por partes, adaptándole poco a poco los sistemas de almacenamiento de datos actuales. No sólo hay que cambiar el medio, sino también el cabezal de grabación y la electrónica de procesamiento de las señales. Aunque algunas compañías han venido estudiando con ahínco, durante una década, la grabación perpendicular, la in-



6. DISPOSITIVOS MAGNETO-OPTICOS: emplean láseres para escribir y leer datos. Para escribir, el láser calienta un punto del medio cuya coercividad disminuye con el aumento de temperatura, haciendo posible magnetizar fácilmente el punto caliente con un débil campo magnético aplicado. Para leer los datos, el haz láser se conmuta a una intensidad más baja y se polariza. Ahora bien, el plano de polarización gira cuando el medio magnetizado refleja la luz polarizada; y así, un segundo filtro polarizador (el analizador) convierte el cambio de polarización del haz reflejado en un cambio de intensidad de luz.

dustria de almacenamiento de datos se ha ido atrincherando, al propio tiempo, en la tecnología de grabación longitudinal. Y puesto que parece posible alcanzar mayores densidades de bits continuando con el avance de la tecnología establecida (sobre todo con el desarrollo de medios de láminas finas), las empresas carecen de estímulos para invertir en la modificación de las tecnologías de grabación.

La grabación longitudinal continuará, así parece, dominando en las unidades de disco duro, al menos. Por otro lado, es improbable que los medios de láminas finas soporten la abrasión que sufrirían en las unidades de cinta y disco flexible, donde el cabezal está en contacto con el medio. En el caso de dispositivos de almacenamiento de contacto, la grabación perpendicular puede incorporarse si se encuentra un medio suficientemente resistente al desgaste.

El rápido progreso experimentado por los dispositivos de almacenamiento de datos y el crecimiento vertiginoso de ese sector industrial ha cerrado las puertas a cualquier otra tecnología de almacenamiento de datos que pretenda afianzarse en un segmento importante del mercado en los próximos cinco años. Pero hay otras tecnologías que podrían entrar en liza. Me refiero, por ejemplo, a la tecnología de grabación óptica, que prescinde de los medios magnéticos y ofrece una densidad lineal de bits más alta y un coste más bajo por bit almacenado que las unidades de disco magnético. Ello no obstante, las unidades ópticas actuales aportan bastante menos prestaciones que las unidades magnéticas en el dominio de las velocidades de grabación y búsqueda de datos y tiempos de acceso. Además, no son borrables: los datos pueden escribirse en un disco óptico hasta su completa ocupación y leerlos después cuantas veces se quiera, pero no pueden escribirse nuevos datos en el lugar de los antiguos.

Las limitaciones de un sistema óptico de almacenamiento de datos se soslayan sin perder sus ventajas recurriendo a la grabación magneto-óptica; se trata de un sistema híbrido que utiliza un haz láser para escribir y leer datos en un medio magnético especial. Igual que sucedía con los medios de grabación perpendicular, se procura que el medio de los discos magneto-ópticos tenga su eje preferido de magnetización perpendicular al plano del disco. Pero uno y otro tipo se distinguen en un elemento clave: el medio magneto-óptico exhibe una alta coercividad a temperatura ambiente y baja coercividad a temperaturas más altas. Así pues, puede enfocarse un haz láser sobre una pequeña región del medio magneto-óptico, para calentarla hasta la temperatura en que la aplicación de un campo magnético provoque la magnetización de la región. Como en los medios de grabación magnética, el patrón de magnetización codifica datos binarios en los medios magneto-ópticos.

Para leer los datos registrados se apela al efecto Kerr. Este fenómeno consiste en una rotación del plano de polarización que un haz de luz sufre cuando es reflejado por un medio magnético. Según la magnetización en el medio magneto-óptico apunte hacia arriba o hacia abajo, la rotación será en el mismo sentido o en sentido contrario de las agujas del reloj. Los datos pueden leerse entonces en el disco magneto-óptico detectando la rotación en la polarización que se encuentra cuando un haz láser polarizado se refleja en el medio. La lectura de datos del disco magneto-óptico puede hacerse con el mismo láser que los escribió, aunque operando a una fracción de la potencia utilizada para escribirlos.

Los medios que mejor se adaptan a la grabación magneto-óptica constan de aleaciones de "tierras raras" (gadolinio y terbio) y metales de transición (hierro y cobalto), depositados por pulverización sobre un disco. Los momentos magnéticos de los elementos de tierras raras apuntan en direcciones opuestas de los momentos de metales de transición, de manera que la magnetización neta en el volumen de la aleación es la diferencia de magnetización entre los dos elementos. Ahora bien, las magnetizaciones de los dos tipos de elementos varían de modo distinto con los cambios de temperatura; por consiguiente, puede lograrse que la magnetización neta y la coercividad de la aleación se modifiquen con la temperatura.

En particular, la magnetización de las tierras raras domina, en la aleación, a las temperaturas más bajas; la magnetización opuesta de los metales de transición domina a temperaturas altas. A cierta temperatura intermedia, en el denominado punto de compensación, la magnetización neta de la aleación es nula. Pero si la temperatura crece bastante, la magnetización neta vuelve a aproximarse a cero, porque la agitación térmica de los átomos adquiere inten-

7. CAMPOS DESMAGNETIZADORES en un medio magneto-óptico. Se aplican a la escritura y borrado de regiones; de ese modo, queda eliminada la necesidad de un campo magnético externo. Un pulso de láser confiere al medio una distribución de temperatura característica (*a*), que reduce la coercividad en una región magnetizada uniformemente (*b*), permitiendo que el campo desmagnetizador interno conmute la magnetización de la región calentada (*c*). Se invierte también el campo desmagnetizador de la región conmutada, de manera que un pulso posterior puede provocar que la magnetización del centro de la región conmutada se invierta a su dirección original (*d*). Las partes externas de la región doblemente conmutada, inestables, se anulan y dejan al medio como estaba en un comienzo (*e*). (Esquemas realizados por Ian Worpole.)

sidad suficiente para convertir en aleatoria la orientación de los momentos magnéticos atómicos. Ese efecto se produce a la temperatura de Néel.

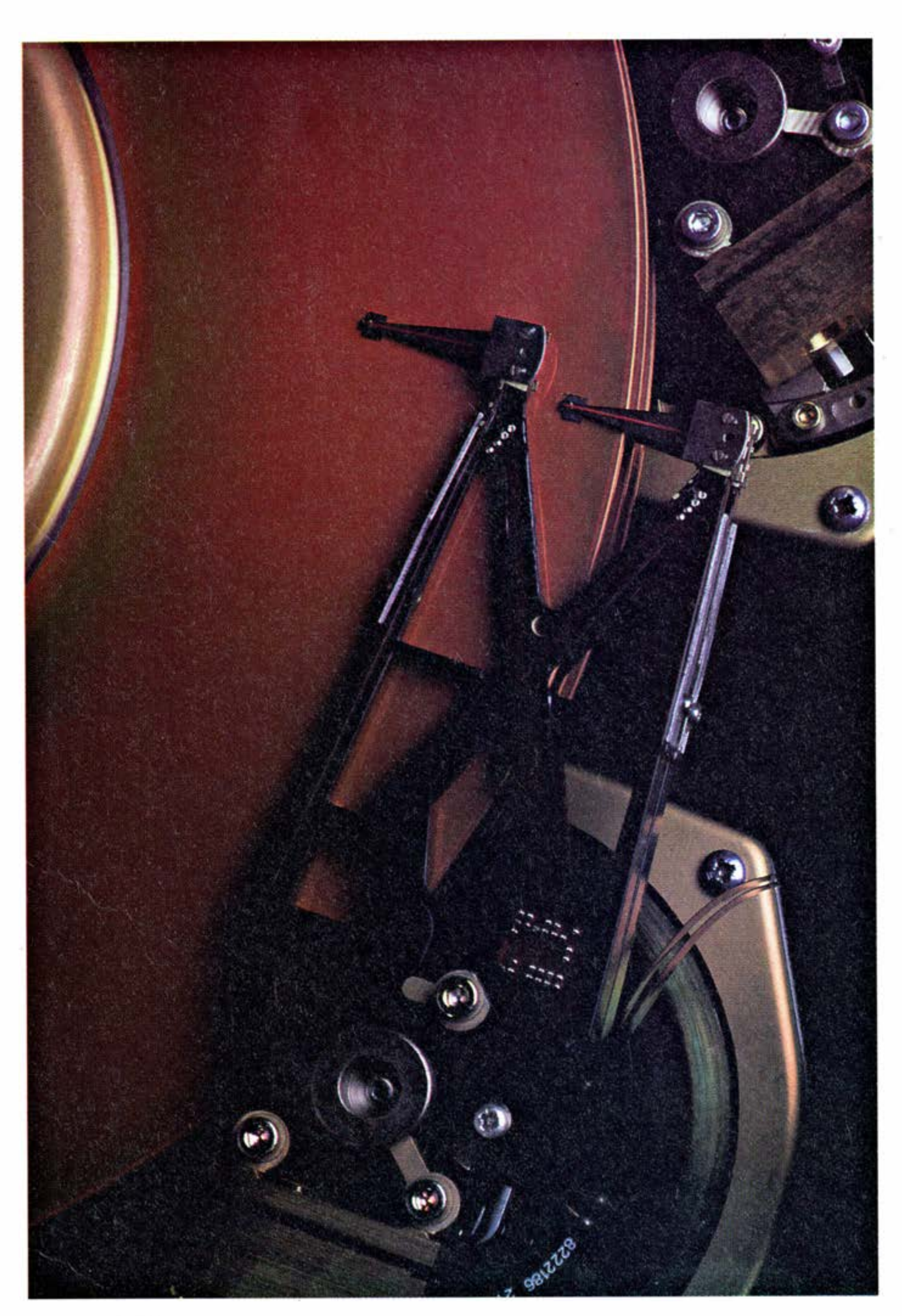
En el punto de compensación, la coercividad tiende a infinito, por la simple razón de que el campo magnético aplicado no tiene magnetización con la que interaccionar. Por contra, debajo mismo de la temperatura de Néel, basta un campo magnético débil para determinar la dirección en que se magnetizan los átomos de los metales de transición. En virtud de ello, los medios magneto-ópticos exhiben una alta coercividad cerca del punto de compensación y una coercividad decreciente conforme se acerca a la temperatura de Néel.

Desde hace tiempo se sabe que los discos magneto-ópticos son, en principio, borrables y reescribibles. Ello no ha impedido que muchos investigadores vinieran creyendo en la imposibilidad de invertir la magnetización en un medio magneto-óptico a la misma velocidad en que podían escribirse los datos sobre un medio virgen. Suponíase que los campos requeridos para invertir la magnetización del medio —incluso cuando éste se halla calentado con láser— eran demasiado intensos. Hubiese sido muy difícil alcanzar un cabezal capaz de generar tales intensidades de campo e invertirlos en coordinación con una velocidad razonable de grabación de datos. Por eso, los investigadores pensaron que, para reescribir un disco magneto-óptico, sería necesario girar el disco una revolución bajo la cabeza óptica que borrara los datos antiguos, y girar otra revolución que escribiera los nuevos, desacoplando la velocidad de borrado de datos de la velocidad de grabación de otros. Tal diseño, sin embargo, añadiría varios milisegundos al tiempo requerido para reescribir sobre un disco.

Se ha demostrado que la escritura directa de un medio magneto-óptico es posible. Sólo hay que intensificar los campos desmagnetizadores en el medio del disco, o rebajar la coercividad a alta temperatura del medio, de suerte que la magnetización conmute siempre que se caliente el medio; no se necesita aplicar ningún campo magnético externo [*véase la figura 7*]. En una posible disposición, un sistema láser de baja potencia lee la dirección de magnetización de una región del disco y se la compara entonces con la dirección que los nuevos datos asignan a la magnetización de la región. Si procede invertir la magnetización, un segundo sistema láser manda un pulso de luz sobre la re-

gión y la calienta, de suerte que los campos desmagnetizadores puedan conmutar la dirección de magnetización. En una vía alternativa, el láser de guía calienta el medio de la región donde va a reescribirse; lo hace con un pulso lasérico de duración tan corta que las regiones con magnetización invertida tornan a su estado original, mientras deja intactas a las demás; asegura, pues, que la magnetización de todas las regiones apunte en la misma dirección. El segundo láser emite después pulsos de luz más largos para conmutar la magnetización de las regiones que deben invertirse de acuerdo con los nuevos datos. En uno y otro caso, podemos enfocar ambos haces de láser a través de una sola lente; sus puntos focales sólo exigen distar unas cuantas micras.

Los láseres de lectura/escritura pueden enfocarse sobre un punto diminuto (del orden de media micra), y seguir una pista con una precisión de 0,1 micra; ello posibilita la existencia de pistas de datos de extrema finura. En este sentido, los sistemas magneto-ópticos experimentales de almacenamiento de datos han demostrado que son posibles densidades de bits por unidad de área superiores, en más de un orden de magnitud, a las de las unidades de disco magnético en uso. Además, el sistema de enfoque de la lente puede mantenerse a varios milímetros del medio, sin perder su elevada resolución a través de una capa transparente. El amplio intervalo entre la lente y el medio nos libra de preocuparnos por los aterrizajes del cabezal; a su vez, la capa transparente protege al medio de la degradación ambiental. Y lo que es más: la materia extraña que aparezca en el sustrato transparente no está en el plano focal de la lente y, por tanto, no produce errores siempre que ese cuerpo no sea muy grande. (Esa es la razón por la que los discos compactos, que también leen datos ópticamente, son móviles.)



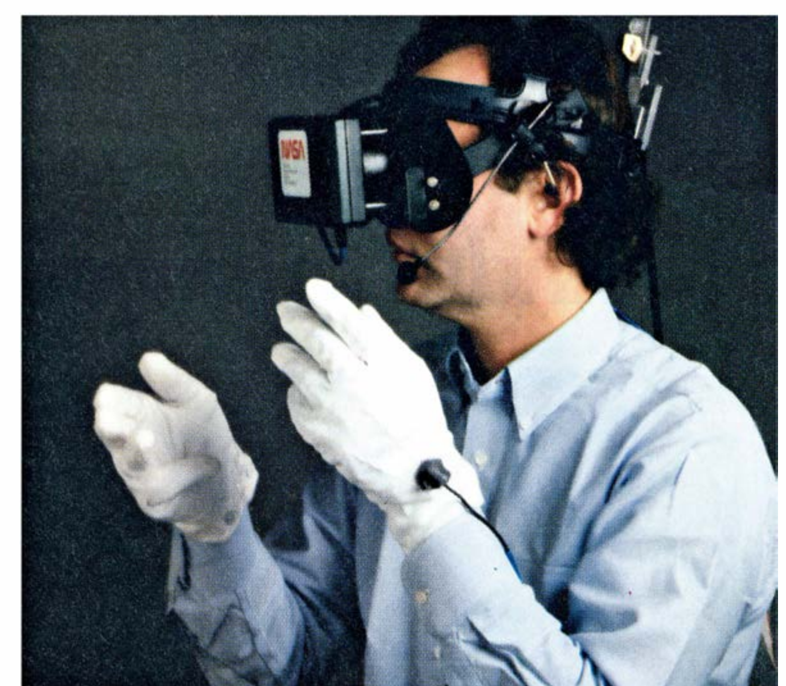
8. UNIDADES DE DISCO MAGNETICO construidas por IBM; utilizan cabezales electromagnéticos de lámina delgada (en los extremos de los brazos) para escribir datos rápidamente y a elevadas densidades.

Además de los componentes ópticos asociados con la lectura y escritura, el cabezal magneto-óptico aloja también componentes que mantienen enfocado el haz láser y posibilitan que la lente siga adecuadamente la pista de datos sobre el disco en rotación. Todas estas piezas están en un cabezal que mide un par de centímetros escasos de lado y pesa unos 150 gramos. Con todo, es muy grande comparado con un cabezal de grabación magnética normal. Debido a la mayor masa de los cabezales ópticos, éstos no pueden mudarse rápidamente de una pista a otra: los tiempos de acceso en unidades ópticas han sido de 0,1 segundos, un orden de magnitud mayor que los tiempos de unidades magnéticas.

Parece probable, sin embargo, que los tiempos de acceso de unidades magneto-ópticas se acortarán bastante. En este sentido, algunos investigadores han demostrado que las fibras ópticas pueden unir las lentes con el resto de componentes ópticos. En ese sistema sólo muda la estructura asociada con la lente desplazada para acceder a una pista del disco, estructura cuyo peso podría acercarse al del cabezal magnético. Otros investigadores proponen recurrir a la microelectrónica para integrar la óptica y los dispositivos electrónicos en un todo complejo. Todo indica, pues, que éstas y otras innovaciones acortarán los tiempos de acceso de las unidades magneto-ópticas, hasta equipararlos a los de las unidades magnéticas.

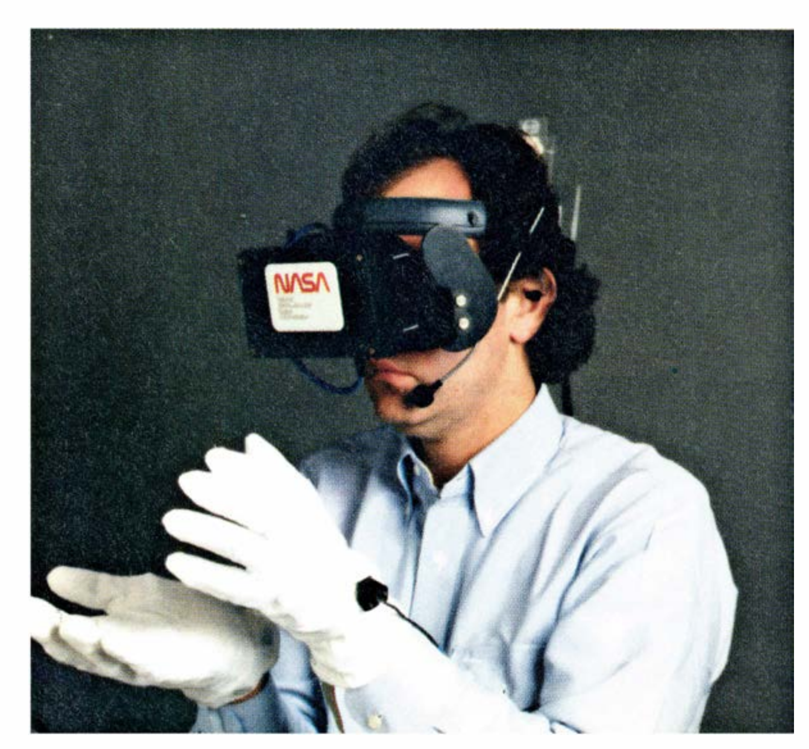
La capacidad de escribir datos nuevos sobre los antiguos, combinada con las altas densidades de almacenamiento de información (del orden de 10 millones de caracteres por centímetro cuadrado), la despreocupación por el aterrizaje de los cabezales y la movilidad de los medios ópticos convierten a la grabación magneto-óptica en una poderosa rival de la grabación magnética al uso en las aplicaciones de almacenamiento de datos.


NASA




NASA


NASA


NASA


NASA

Interfases para ordenadores avanzados

¿Por qué tiene que ser difícil utilizar los ordenadores complejos? La próxima generación de superordenadores creará refinadas "entidades artificiales" para una comunicación fluida entre la máquina y el usuario

James D. Foley

El simulador de vuelo constituye un buen ejemplo de la capacidad de los ordenadores modernos para imitar la realidad. El computador combina sonido, fuerza y movimiento que semejan el comportamiento aerodinámico de un avión; los superordenadores especializados aportan las imágenes. Esas representaciones visuales, eficacísimas, no son fáciles de crear; pero los superordenadores pueden convertirse en maestros de la ilusión, un poder que, paradójicamente, apenas si lo saben aprovechar las aplicaciones científicas de nuestros días.

¿Por qué limitar las capacidades de simulación de los ordenadores a una cabina de mandos? ¿No podría la máquina, facultada para recrear las sensaciones de vuelo, sintetizar los contextos en que se plantean los problemas científicos? ¿No podría construir un entorno comunicativo más natural que el teclado y las órdenes escritas? En definitiva, ¿no sería posible programar un ordenador para que creara "entidades artificiales" a través de las cuales comunicarse con el usuario?

Para muchos ingenieros y profesionales de la informática la respuesta a esa pregunta es un rotundo sí. Se han elaborado ya anteproyectos de entidades artificiales que revisten incluso mayor complejidad que las simulaciones de vuelo. Se están desarrollando técnicas de interfases que adaptarán mejor los superordenadores a las formas humanas de comunicación: táctil, gestual, hablada y visual. Además de visualizadores gráficos más realistas, la próxima generación de superordenadores ofrecerá la manipulación manual real de las imágenes generadas por la máquina, junto con sensaciones táctiles y contrarreacción de fuerza ("force feedback"). Los sensores medirán la posición de la cabeza y seguirán el movimiento de sus ojos; los programas de reconocimiento de la voz permitirán a los ordenadores interpretar el lenguaje hablado.

Los investigadores esperan que el aprendizaje y la explotación de las capacidades del superordenador resulten más eficaces y entretenidas a través de esas entidades artificiales. Tales interfases, muy complejas, podrán realizar para los científicos y técnicos la función que han desempeñado los programas de hojas electrónicas de cálculo como el Lotus 1-2-3 para los contables. Ni que decir tiene que esas "realidades artificiales" aparecen más prometedoras para ciertos problemas que para otros; es difícil imaginar, por ejemplo, cómo se beneficiaría una autoedición de la representación tridimensional y de las pantallas del tamaño de una pared. En el otro extremo, muchos problemas científicos, especialmente los que pueden representarse en tres dimensiones, exigen un mayor grado de interacción entre hombre y máquina.

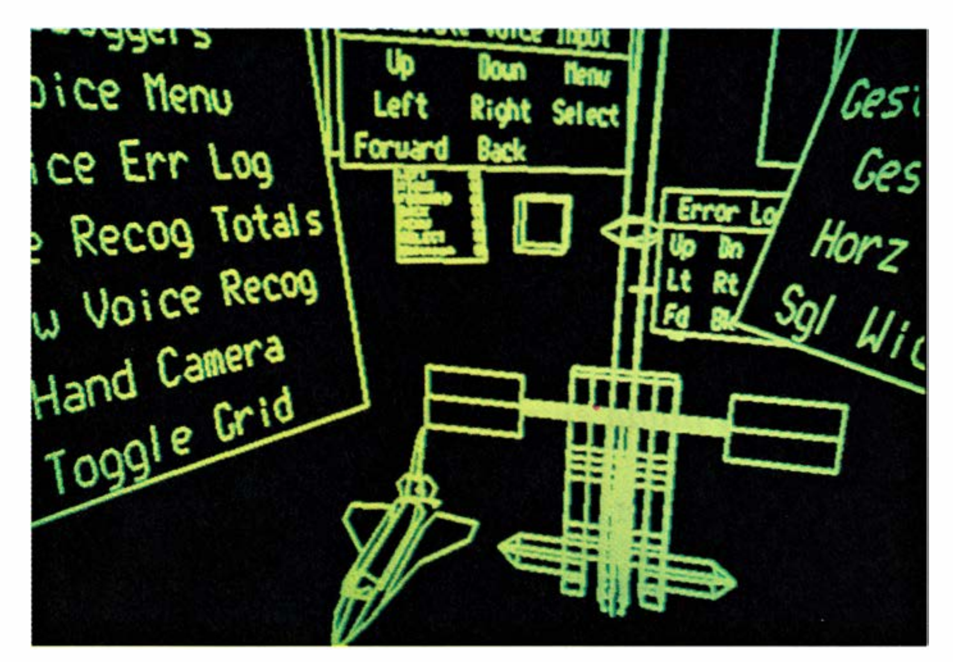
La interfase entre el usuario y el ordenador puede constituir el último peldaño del diseño de ordenadores. En los últimos años, los costes del soporte físico ("hardware") han caído drásticamente; también están bajando, aunque más lentamente, los costes del soporte lógico ("software"). Para mejorar el rendimiento de los ordenadores y minimizar el uso de memoria contamos ya con ciertas técnicas bien asentadas, en constante refinamiento. La cuestión pendiente es que el usuario saque el máximo rendimiento.

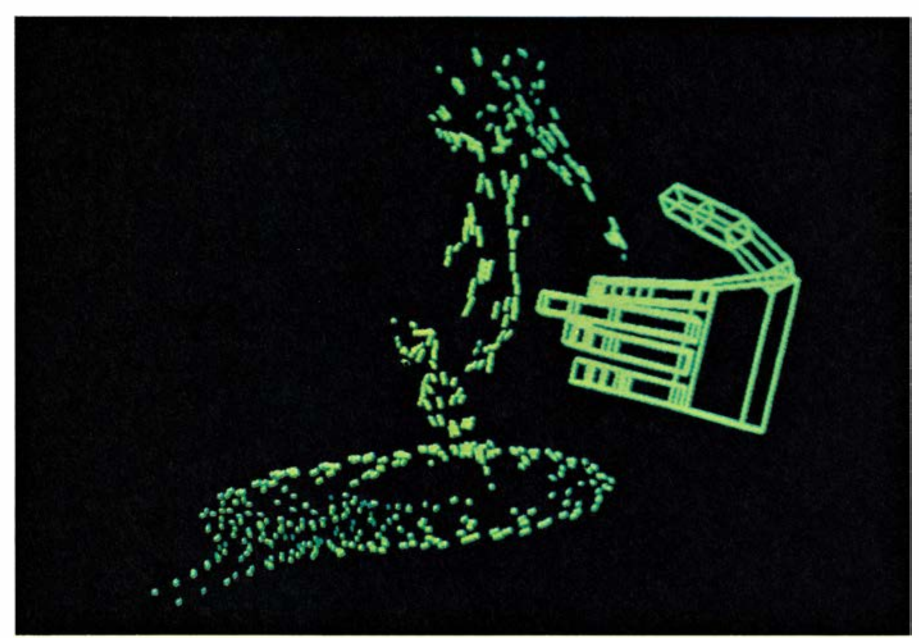
El cálculo científico se ha visto obligado a concentrarse en las prestaciones del soporte físico en razón de sus propias necesidades de cómputo, ingentes. La próxima generación de superordenadores posibilitará, a buen seguro, que ese interés cambie por fin. El Cray-2, uno de los ordenadores más rápidos del mundo, puede acometer 1000 millones de operaciones de coma flotante por segundo (MFLOP). (FLOP es la unidad industrial estándar de la velocidad de cálculo; los ordenadores personales funcionan a velocidades de 1000 a 100.000 FLOP.) La próxima generación de Cray, proyectada para el año que viene, decuplicará ese ritmo. Las mismas estaciones de trabajo que suministran los gráficos de los superordenadores están adquiriendo mayor potencia: la estación de trabajo que Stellar Computer Inc. espera comercializar en enero funcionará en el rango de los cincuenta MFLOP.

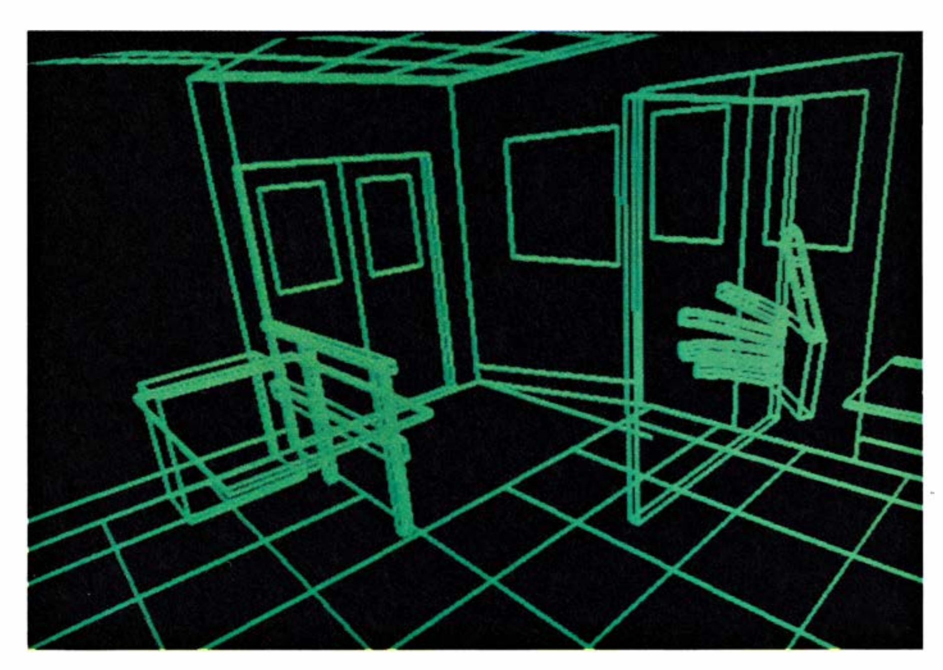
¿Cómo conseguir que esta potencia adicional logre que la interacción entre el usuario y el ordenador se convierta en una cooperación y pierda el aire de enfrentamiento que tiene hoy? Una mayor celeridad en las operaciones de cálculo constituye, de suyo, un avance, disminuyendo, y en algún caso eliminando, el período de espera durante el cual el ordenador genera los resultados. Eso es obvio. No lo son tanto los medios para integrar más estrechamente a los usuarios del superordenador en el proceso de resolución de los problemas.

Utilizar hoy un superordenador vie-

1. MONITOR CON FORMA DE VISOR dotado de un sensor de posición y orientación, guantes que siguen los movimientos de la mano y de los dedos y un micrófono conectado para el reconocimiento de la voz: esos elementos llevan al usuario hasta una entidad creada por el ordenador. El usuario da instrucciones al ordenador señalando, hablando, gesticulando y manipulando imágenes gráficas. Los investigadores del Centro Ames de la NASA han construido distintas entidades artificiales para su empleo en este sistema.







2. PROYECCION DE GRAFICOS en el interior del visor ofrecido en la figura 1. Entre ellos, vemos menús de interfases (*arriba*), modelos de corrientes de aire (*centro*) y el laboratorio que alberga al sistema (*abajo*). Las fotografías dan sólo una idea del realismo de los visualizadores; éstos suministran indicios de profundidad mostrando a cada ojo una vista desde una perspectiva diferente y permiten al usuario recorrer el entorno generado por el ordenador con sólo girar la cabeza. El usuario puede seleccionar una opción del menú con una palabra o un gesto, rotar un modelo de corriente de aire para contemplarlo desde otro ángulo o alargar su mano y "tocar" las paredes y bancos del laboratorio.

ne a ser como consultar una oráculo de la era de las máquinas. Hay un período de preparación, durante el cual se define un problema y se especifican sus parámetros en una estación de trabajo. La estación organiza el problema en una forma apropiada para el superordenador, pudiendo tardar segundos, horas o días en los correspondientes cálculos. Estos no pueden interrumpirse o alterarse por el usuario, una vez que han comenzado; si los resultados sugieren la comprobación de parámetros alternativos, debe repetirse la operación desde el principio.

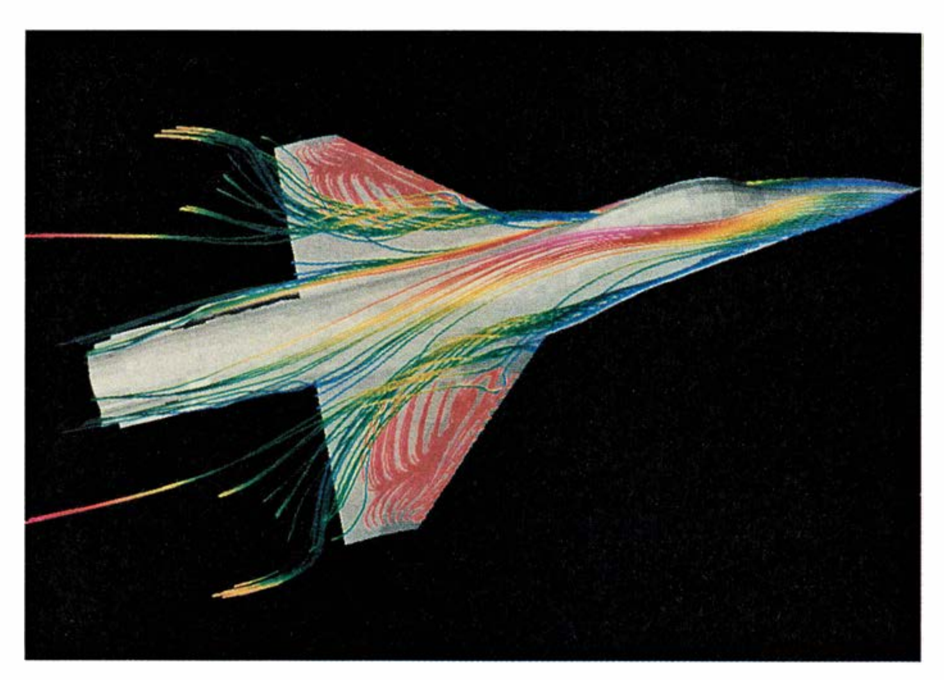
La llegada de las creaciones artificiales, anunciada por la simulación de vuelo, cambiará de raíz el trato del hombre con el superordenador. Esas entidades permitirán al usuario relacionarse con el ordenador en un formato intuitivo y directo y aumentar el número de operaciones por unidad de tiempo. La investigación en el campo de las creaciones artificiales se propone como objetivo final desarrollar un entorno simulado que parezca tan "real" como la realidad que remeda. El máximo potencial de las interfases podría residir, sin embargo, en su capacidad de trascender la propia realidad, modelando de manera concreta entidades abstractas (ecuaciones matemáticas) y permitiendo a los usuarios superar problemas de escala mediante la manipulación, por igual, de átomos y galaxias.

Las creaciones artificiales constan de tres componentes: imágenes, comportamiento y relación. Las imágenes visuales realistas ayudan al usuario a interpretar la información que le presenta el ordenador. Pueden ilustrar objetos reales, como estructuras de construcción, o abstractos, como modelos del flujo de un fluido. Las imágenes se comportan igual que lo harían los objetos o las abstracciones que representan. Modelar un determinado comportamiento exige una mayor cantidad de cálculo, porque a menudo supone la resolución iterativa de extensos sistemas de ecuaciones. Finalmente, el usuario

se relaciona con una entidad artificial de acuerdo con las mismas pautas que sigue con el mundo tridimensional: moviendo, señalando y recogiendo cosas, hablando y observando desde varios ángulos diferentes.

El desarrollo del componente interactivo o relacional de las entidades o creaciones artificiales se halla retrasado respecto de los otros componentes. Las técnicas comerciales para observar y manipular en tres dimensiones apenas si han empezado a andar. ¿Qué tipos de dispositivos de interacción restablecerán el equilibrio?

Comenzaré por el componente más familiar de la interfase de un ordenador: el monitor de visualización. La estación de trabajo normal de un ordenador está equipada con un monitor de color de 19 pulgadas de diagonal; observado desde una distancia de 60 centímetros, el visualizador subtiende 37 grados de campo visual horizontalmente y 28 grados verticalmente. Con todo ello, el campo visual de un ojo abarca 180 grados horizontalmente y 150 grados verticalmente, suponiendo que la cabeza esté fija. Los visualizadores que abarcan todo el campo visual confieren al observador la sensación de formar parte de la escena, en lugar de estar mirando desde fuera; así lo corroboran los sistemas de cine en pantalla gigante IMAX y OMNIMAX. Por tanto, las proyecciones del tamaño de una pared de las pantallas de ordenadores podrían sumergir al usuario en el propio mundo de la entidad artificial creada.



3. **SIMULACION DE MODELOS DE CORRIENTES DE AIRE producidas por un caza F-16A. Constituye un ejemplo de la perfección alcanzada por los gráficos de los superordenadores. El recorrido del aire y su elevación por encima del cuerpo del avión se indican con líneas de colores: el azul significa elevaciones bajas y el rojo altas. Las turbulencias rojas de las alas descubren zonas de esfuerzo de cizallamiento. La información que esta imagen aporta, ayuda a los ingenieros a diseñar aviones que ofrezcan una menor resistencia aerodinámica y mayor sustentación. Gráficos de parejo refinamiento se logran ya de manera rutinaria.**

Un dispositivo visualizador mayor abarcaría mejor el campo visual, pero no mostraría más detalles si utilizara el mismo número de pixeles o elementos de imagen. Los pixeles son los puntos de luz discretos que forman una imagen en un tubo de rayos catódicos y definen la resolución de ésta. El monitor de una estación de trabajo consta, por lo común, de un millón de pixeles dispuestos en una red de 1280 por 1024 unidades. A una distancia normal de observación, cada píxel subtiende alrededor de dos minutos del campo visual, si bien el ojo humano puede distinguir detalles de hasta un minuto. Existen monitores de color de 20 pulgadas cuadradas con unos cuatro millones de pixeles en una red de 2000 por 2000, en los que cada píxel subtiende un ángulo de aproximadamente 1,4 minutos observado desde una distancia de 60 centímetros. Todavía está por construirse un monitor que aproveche plenamente la agudeza de la vista humana a una distancia normal de observación.

Dadas estas limitaciones, ¿cómo puede mejorarse todavía más el realismo? Una alternativa es el visualizador en forma de visor. Este visualizador puede facilitar la percepción de profundidad de la misma manera que un Vu-Master logra su efecto estereoscópico: cada ojo percibe una perspectiva ligeramente compensada de la misma imagen. Otras indicaciones de profundidad provienen de la paralaje de movimiento, que es el fenómeno que describe el desplazamiento del fondo que ocurre cuando un observador —que tiene la mirada puesta en un punto del espacio— cambia de posición. Este efecto se logra con la ayuda de un sensor, que registra la posición y la orientación de la cabeza. El sensor reconoce movimientos drásticos de la cabeza; y así, al girar ésta, el usuario puede disfrutar de la ilusión de explorar un panorama artificial. Las imágenes que se ven dependen de la dirección en que se orienta el usuario.

El primer visualizador en forma de visor lo construyó Ivan E. Sutherland en la década de los sesenta. En él, tubos de rayos catódicos en miniatura actuaban como visualizadores y se comunicaba la posición y la orientación de la cabeza al ordenador mediante uniones mecánicas. Los actuales monitores ligeros de cristal líquido y sensores electrónicos han simplificado su misión. El sistema más avanzado que incorpora estos adelantos no sólo crea entidades artificiales, sino que reemplaza también una creación por otra. En el Centro de Investigación Ames, perteneciente a la NASA, Scott S. Fisher, Michael W. McGreevy y James C. Humphries han construido un casco para su uso en una estación espacial que proyectaría, a un astronauta situado en el interior de la estación, lo que "viera" un robot que operase en el exterior. Cuando el astronauta gira la cabeza, los ojos de la cámara del robot se mueven en la misma dirección.

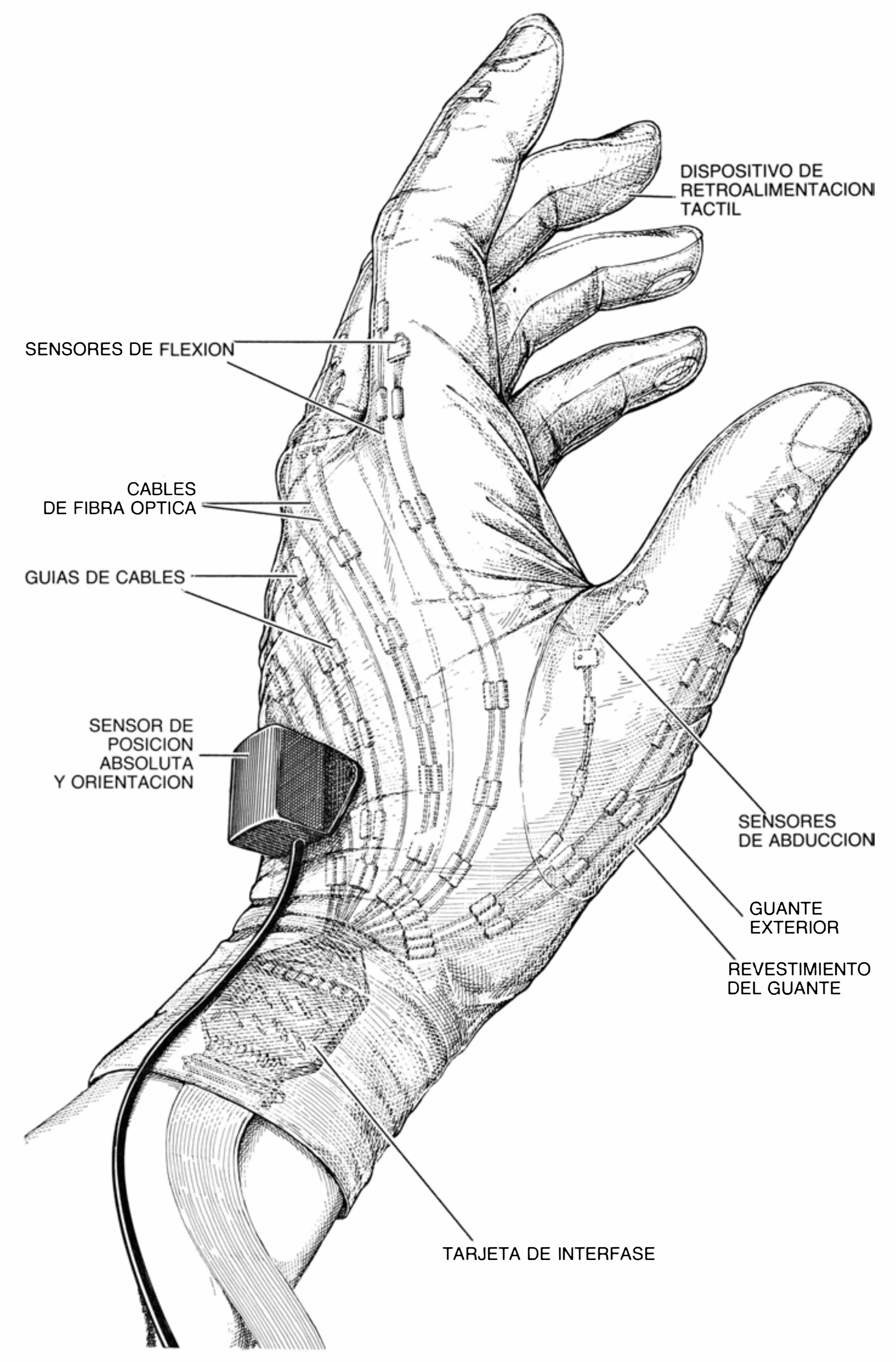
El sensor electrónico que registra la posición y la orientación de la cabeza en el sistema de la NASA constituye una parte, muy importante por cierto, de otros muchos sistemas de interfases. Fabricado por la división de Ciencias de Navegación Polhemus de McDonnell Douglas Corporation, el sensor envía pulsos electromagnéticos desde tres antenas transmisoras hacia tres antenas receptoras. En ambas unidades, la transmisora y la receptora, las bobinas de la antena se hallan en ángulo recto y forman un sistema de coordenadas cartesianas.

El transmisor es una caja de unos cinco centímetros de lado que debe colocarse a un metro y medio del receptor. Emite una secuencia de tres pulsos, uno desde cada antena. Los pulsos inducen una corriente en las bobinas del receptor, que es un cubo de dos centímetros de arista colocado en el objeto que se rastrea. La magnitud de la

corriente depende de la distancia entre el receptor y el transmisor, así como de la orientación relativa de las bobinas de ambos. Un ordenador puede calcular la posición tridimensional de la unidad receptora a partir de los nueve valores de corriente que se obtienen de tres pulsos sucesivos. Los tres pulsos se repiten unas 40 veces por segundo; las imágenes resultantes se mueven algo erráticamente. No habrá, a buen seguro, un movimiento simulado uniforme mientras el sensor no pueda producir 60 tripletes de pulsos por segundo.

Instalado en el casco de la NASA, el sensor Polhemus sería bastante diestro en la determinación de la dirección de la mirada del usuario. Menos en un detalle: los ojos pueden moverse, y lo hacen a menudo, independientemente de la cabeza. Para superar esta dificultad, los ingenieros están estudiando una técnica habitual en psicología experimental. Desde hace tiempo, los psicólogos han venido utilizando los rastreadores de ojos para recoger datos sobre el modo como la gente lee y contempla grabados o imágenes. Los rastreadores de ojos hacen rebotar un haz de luz en la córnea del ojo. La dirección en que se refleja la luz indica dónde está mirando el usuario: su punto de atención.

Los rastreadores de ojos acaban de introducirse en el campo de los ordenadores. Los que se acoplan a las gafas cuestan unos miles de dólares, pero no son muy precisos. Cierto sistema más complicado proyecta un punto de luz infrarroja sobre la córnea y detecta su reflexión con una cámara de televisión gran angular colocada a un metro del usuario. La cámara permanece fija sobre el ojo aunque se produzca un profundo movimiento de cabeza, a menos que el movimiento sea brusco; podría ser perturbada, por ejemplo, por un estornudo.



4. DATAGLOVE desarrollado por VPL Research Inc.; traduce movimientos de la mano y de los dedos en señales eléctricas. Entre dos capas de tela, los cables de fibra óptica fijados en ambos extremos a una tarjeta de interfase recorren la longitud de cada dedo por partida doble. Cada cable tiene un diodo emisor de luz en un extremo y un fototransistor en el otro. La disposición de los cables determina que la luz escape cuando se dobla un dedo; el fototransistor convierte la luz que recibe en una señal eléctrica. El sensor de posición y orientación está fabricado por la división de Ciencias de Navegación Polhemus de la McDonnell Douglas.

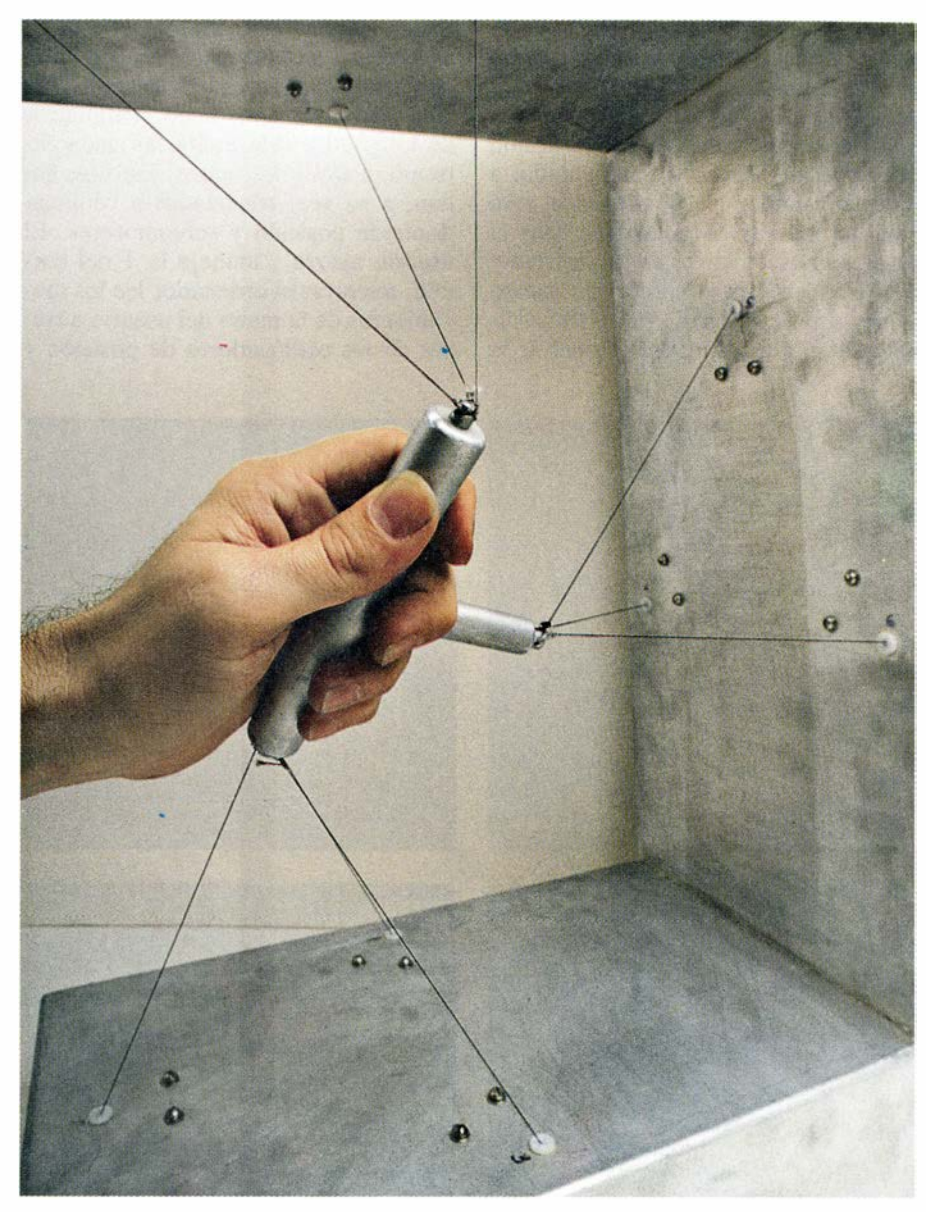
Pantallas del tamaño de una pared, visualizadores en forma de visor, sensores de posición y rastreadores de ojos pueden mejorar la credibilidad de una creación artificial por ampliación del campo visual y por la perfección de los detalles en el punto de atención. Los visualizadores pueden presentar imágenes reales, como lo hace el casco para estación espacial de la NASA, o imágenes artificiales generadas por un superordenador. ¿Cómo se las habrá el usuario con estos visualizadores?

La mayoría de los dispositivos actuales de interacción se limitan a dos dimensiones. Las propias manipulaciones de las simulaciones tridimensionales de un ordenador deben realizarse a través de un medio bidimensional, un ratón o una palanca de mando (“joystick”). Mayores obstáculos oponen los selectores de posición, que controlan los tres ejes por separado. Lo ideal sería que pudiera fabricarse un dispositivo de interacción que combinara la precisión, el control y la agilidad de la mano humana. Pero ocurre que el “dispositivo” existe ya; se trata de la propia mano equipada con un sensor Polhemus y un guante especial que registra los movimientos de la mano y de los dedos.

El DataGlove (“guante de datos”) lo han desarrollado Thomas G. Zimmerman y L. Young Harvill en VPL Research Inc., una empresa californiana. Cables de fibra óptica insertados entre

dos capas de tela recorren los dedos, pulgar incluido. Ambos extremos de cada cable están conectados en una tarjeta de interfase cerca de la muñeca. Un diodo emisor de luz (LED) situado en un extremo manda luz a través del cable hasta un fototransistor, instalado en el otro extremo. El fototransistor convierte la luz en una señal eléctrica, reconocible por el ordenador. La señal viaja desde la muñeca hasta el terminal a través de un cable eléctrico. Aunque los cables normales de fibra óptica transmiten luz cuando se doblan, los del DataGlove están dispuestos en los puntos donde se doblan los dedos, de manera que la luz escapa cuando se tuerce un dedo, o cuando, por ejemplo, el usuario mueve su pulgar hacia su dedo índice. Cuanto mayor sea el movimiento, más luz se perderá.

Acoplado con un sensor Polhemus, que puede instalarse en el dorso de la mano, el DataGlove presenta unas posibilidades fascinantes. El grupo de Fisher en Ames utiliza el guante juntamente con el casco de la estación espacial; la NASA espera que un robot alojado en el exterior de la estación sea, algún día, capaz de llevar a cabo reparaciones y maniobras complejas por imitación de los movimientos de la mano de un astronauta que esté en el interior. VPL también está incorporando los principios del DataGlove en un DataSuit que cubra la totalidad del cuerpo.



5. CORDON DE MANDO. Se trata de uno de los dispositivos de contrarreacción de fuerza más eficaces. Diseñado por Richard J. Feldmann, del Instituto Nacional norteamericano de la Salud, el ingenio comunica a un ordenador la posición de una mano que agarra la T suspendida; el superordenador dirige unos servomotores que ejercen una fuerza a través de la tensión diferencial en los nuevos cables conectados a la T.

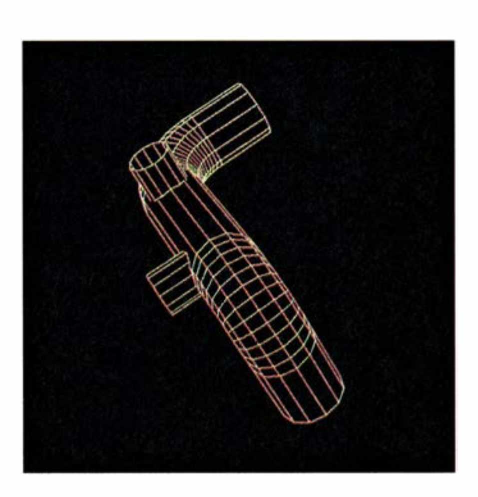
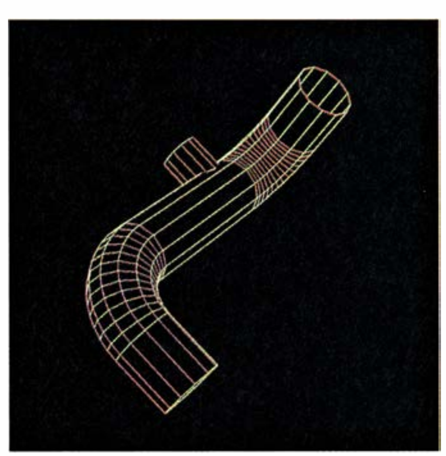
Una interfase tan compleja como el DataGlove quizá no sea la panacea universal, pero muchas de las ecuaciones científicas que se introducen en los superordenadores representan sistemas que admiten mejor un tratamiento directo que a través de órdenes escritas. Además, las entidades artificiales pueden añadir un elemento táctil en los sistemas que escapan a toda posible manipulación. Imaginemos, por ejemplo, un bioquímico que se halla estudiando dos moléculas: una enzima y el sustrato al que se une. Las estructuras de la enzima y del sustrato, conocidas, se representan en el monitor del ordenador. El químico desea averiguar qué parte de la enzima reacciona con la correspondiente del sustrato. Provisto de un DataGlove podría manipular ambas moléculas, como piezas de un rompecabezas, para ver qué partes se acoplan.

Imaginemos, ahora, que se puede sentir la topografía de la molécula de la enzima: sus grietas y sus salientes, sus bordes uniformes y las esquinas agudas. Supongamos que se inspecciona con un dedo el lugar activo de la enzima que ejerce una fuerte atracción química sobre el sustrato. Y hácese lo propio con el sustrato cerca de la zona activa hasta sentir el tirón de las fuerzas interatómicas que unen ambas moléculas. Para lograr estos efectos los ingenieros están investigando métodos capaces de crear una contrarreacción ("feedback") táctil y de fuerza.

En el diseño del DataGlove podrían incorporarse tres innovaciones que se están ensayando para la contrarreacción táctil. Una de ellas está adaptada de un dispositivo de contrarreacción táctil para los invidentes; en éste, pequeños solenoides presionan, contra la piel, unos cables romos. Quizás ese tipo de accionador resulte, sin embargo, demasiado grande: cada unidad viene a medir unos ocho milímetros de espesor. También pueden utilizarse, junto con el guante, cristales piezoeléctricos, que vibran cuando se activan por una corriente eléctrica; la mente interpreta su vibración como una presión. Un tercer método aprovecha los nuevos "metales de memoria" que cambian su forma con la temperatura. Pequeños trozos aislados de estos metales podrían orientarse para presionar contra la piel cuando se calientan por una corriente eléctrica.

En un dispositivo en forma de guante es más difícil transmitir la sensación de fuerza que la sensación táctil, aunque los metales de memoria aportan cierta

esperanza en este campo. Hace ya algún tiempo, en 1968, un grupo dirigido por Frederick P. Brooks Jr., de la Universidad de Carolina del Norte en Chapel Hill, adaptó para la contrarreacción de fuerza un dispositivo manipulador a distancia similar a los utilizados para manejar materiales radiactivos. Hoy el sistema más efectivo de contrarreacción de fuerza es el cordón de mando ("joystring"), diseñado por Richard J. Feldmann, del Instituto Nacional de la Salud. Bautizado así en memoria de su antecesor ("la palanca de mando"), el cordón de mando es una sencilla T rígida que mide unos ocho centímetros de longitud y está unida, en cada extremo, a tres hilos tensos. Estos se hallan, a su vez, conectados a codificadores de posición y servomotores. El usuario agarra y maneja la T del cordón, mientras el ordenador lee los movimientos de la mano del usuario a través de los codificadores de posición y genera una contrarreacción de fuerza y de par mediante unos servomotores.





6. TUBOS GIRADOS EN DIFERENTES POSICIONES que auxiliaron al autor en su estudio de la incidencia del realismo. A los sujetos de la prueba se les mostraban dos imágenes, una al lado de otra, y se les preguntaba si una era o no la rotación de la otra. El ensayo puso de manifiesto que los usuarios comparaban los tubos coloreados y resaltados (*centro*) un 20 por ciento antes que lo conseguían con las figuras definidas sólo por su contorno (*arriba*); el refinamiento posterior (*abajo*) no mejoró significativamente la comparación.

Aparte del cordón de mando, sólo se han construido unos cuantos dispositivos de contrarreacción de fuerza, cada uno de ellos como pieza única. Los sistemas se han limitado a aplicaciones especializadas en laboratorios de investigación. Los diseños más sencillos pueden responder a la aplicación de una fuerza, pero no pueden generar una contrarreacción. En la Universidad de George Washington mi colega John L. Sibert y su grupo han desarrollado una nueva aplicación de ese sistema de inserción de fuerza: un sistema de pintura para artistas que está basado en una tablilla de datos y percibe, no sólo la posición de un puntero, sino también la orientación de éste y la fuerza que se le aplica. El ordenador emplea esta información para simular el comportamiento de un pincel. Al aumentar la presión hacia abajo, el "pincel" se extiende y la anchura de la línea que se dibuja cambia con la orientación del puntero. La entidad artificial de Sibert refleja hasta cierto punto los útiles tradicionales del artista, pero crea también herramientas artísticas que no existían hasta ahora.

¿De qué otra forma desearía el usuario relacionarse con el superordenador? En muchos casos, hablar o gesticularle a la máquina puede resultar más apropiado o conveniente que recurrir a las manipulaciones no verbales de imágenes simbólicas. Ambos sistemas de reconocimiento de voz y de gesto están más avanzados que los demás componentes de las entidades artificiales. Hace más de 10 años que Nicholas Negroponte y Richard A. Bolt, del Instituto de Tecnología de Massachussetts, demostraron la viabilidad del reconocimiento de la voz en la relación con el ordenador. Funcionan ya máquinas con vocabularios de varios cientos de palabras en aquellos trabajos que requieren el uso de ambas manos; por ejemplo, en la comprobación de montajes electrónicos.

Pero la tecnología para el reconocimiento de la voz no se ha incorporado todavía en la mayoría de los entornos de ordenadores que resuelven problemas. Esta desafortunada tardanza en adoptar una tecnología nueva no es cosa de hoy: el ratón, desarrollado en las postrimerías de los 60, no se comercializó hasta principios de los 80. Mientras tanto, los sistemas de reconocimiento de la voz han ido ganando en refinamiento. Kenneth Davies, del

Centro de Investigación Thomas J. Watson de la IBM Corporation, me mostró recientemente un sistema experimental con un vocabulario de 20.000 palabras: alrededor del 98 por ciento del vocabulario del inglés hablado. La máquina puede incluso interpretar frases fonéticamente enrevesadas como "Write Mrs. Wright a letter right away" ("Escribe inmediatamente una carta a la señora Wright"). IBM espera comercializar su sistema en los próximos años.

La tecnología para el reconocimiento del gesto ha estado languideciendo durante varios años a la espera de la resolución de ciertas dificultades. Hay que enseñar a los sistemas cómo distinguir entre las ambigüedades de una serie de gestos y saber cuándo termina una orden y comienza la siguiente. James R. Rhyne, del Centro de Investigación Watson, ha desarrollado un sistema que reconoce gestos de la mano realizados en dos dimensiones con un dispositivo de posición semejante a un bolígrafo. En una aplicación de hoja electrónica, el usuario puede sumar dos grupos de números encerrando en un círculo cada grupo de una tablilla de datos y haciendo el signo de sumar. El ordenador introduce el total en la posición de la hoja electrónica indicada por el signo de sumar. Aunque esta tecnología está todavía en sus comienzos puedo imaginar el día en que un bioquímico provisto de DataGloves gesticule ante las estructuras moleculares representadas en la pantalla del superordenador y diga: "La fenilalanina 221 de esta hélice (señalando) no interacciona adecuadamente con esta glutamina 57 (perfilando un círculo). Cámbiala por una histidina".

Las creaciones artificiales se encuentran en el buen camino. De hecho, uno de los obstáculos más difíciles que quedan por vencer implica la combinación de técnicas ya en uso. Una vez superada esa dificultad, los dispositivos no requerirán especial dispendio. Un DataGlove, por ejemplo, viene a costar unos 8000 dólares. De hecho el esfuerzo de investigación y programación empeñado actualmente en interfases avanzadas será el aspecto más caro de su realización. ¿Justificarán su coste los beneficios que se obtengan de las entidades artificiales?

A primera vista, diríase que un sistema que invita a la interacción del usuario y que presenta la información en formatos accesibles sería más rápido y fácil de aprender que las interfases convencionales. Se trata de una clase de conjetura muy difícil de cuantificar. ¿Comprendería mejor el ingeniero un análisis de tensiones si pudiera aplicar la tensión con su propia mano? ¿Se conocerían mejor las interacciones moleculares si pudiéramos "sentir" sus fuerzas? ¿Cuán reales han de ser las entidades artificiales para lograr su propósito?

Hace algún tiempo me planteé esta última cuestión a propósito de las imágenes visuales. Trabajando con mi colega Woodrow Barfield y un equipo de alumnos, dirigidos por James W. Sanford, examiné experimentalmente el efecto del realismo sobre la velocidad con que los usuarios desarrollan una sencilla manipulación mental de dos imágenes. La tarea, llamada rotación mental, constituye una prueba habitual en psicología, y a ella se recurre cuando se estudia el proceso mental de representación de imágenes. Mostré a los sujetos dos imágenes de estructuras en forma de tubos, una al lado de otra y ambas diferentes. Se les solicitó que decidieran si una imagen era o no la rotación de la otra. En unas pruebas, las imágenes dibujaban sólo un contorno; en otras se coloreaban y resaltaban, y en unas terceras las imágenes coloreadas y destacadas presentaban superficies con sombreado uniforme. El resultado fue que los usuarios podían comparar las imágenes coloreadas alrededor de un 20 por ciento antes que las figuras definidas por contornos; las representaciones sombreadas más refinadas no mejoraron, sin embargo, el tiempo de comparación. En resumen, la "realidad" de las entidades artificiales puede perderse después de cierto límite.

Desconocemos qué respuesta dar a las cuestiones sobre el valor final de creaciones artificiales. Hace unos 20 años, James J. Batter, alumno entonces de Brooks, observó que quienes estudiaban gráficos de campos de fuerza bidimensionales comprendían mejor los conceptos implicados si no sólo veían los vectores de fuerza, sino que también los "sentían". El estudio de Batter, que utilizó un sencillo dispositivo bidimensional de contrarreacción de fuerza, es el ejemplo más reciente de investigación que corrobora el valor de las entidades artificiales. Con mis experimentos en la Universidad George Washington, ilustra el tipo de investigación necesaria para determinar si el futuro de las creaciones artificiales encierra algo más que mera ilusión.

Avances en redes de ordenadores

Los ordenadores enlazados a través de redes pueden interactuar mutuamente y compartir programas, datos y hasta soporte físico. Pero no basta con el mero enlace material: un ordenador debe entender lo que otro le está diciendo

Robert E. Kahn

¿Por qué los ordenadores se enlazan a través de redes? ¿Por qué se hablan las personas? Las respuestas son similares: para compartir información, acometer acciones o provocar sucesos; en general, para algún tipo de colaboración. Se ha escrito mucho sobre los esfuerzos empeñados en crear ordenadores inteligentes que razonen como los humanos; pero el comportamiento inteligente no se cierra en el ámbito del sujeto. Se basa, en gran medida, en la comunicación: la trama de conexiones, perdurables o transitorias, que el hombre mantiene con sus congéneres. La teoría de la interconexión entre ordenadores es la disciplina que pretende conseguir que los ordenadores se comuniquen, lo que constituye un reto no menor que lograr que razonen.

El problema de la interconexión presenta dos aspectos esenciales. Primero, establecer un enlace electrónico entre ordenadores, amén de un proceso en virtud del cual se transfiera información a través del enlace de una manera secuencial, intacta y libre de errores. Segundo, instituir un lenguaje común, de suerte que un ordenador comprenda lo que el otro le está diciendo a través del enlace.

Abordaré aquí ambos temas. Pero antes es menester dejar clara una idea: hay muchas clases de redes. Los sistemas de reserva automatizados que utilizan las compañías aéreas son de un tipo; en este sentido, el sistema SABRE introducido por American Airlines constituyó uno de los primeros ejemplos de comunicación entre ordenadores a través de líneas telefónicas. Otro tipo de red es el sistema de alerta precoz para la defensa aérea. Se trata, en ambos casos, de sistemas de propósito único, que comportan una escasa interacción directa entre usuarios. (En el sistema de reservas hay una interacción indirecta determinada por el espacio disponible, finito, en cada avión.) Las redes de las que me ocuparé son aquellas que admiten una interacción, directa y rápida, entre usuarios y máquinas, con distintos propósitos. Mas, a pesar de que se ha avanzado mucho en la última década en el desarrollo de tales sistemas, queda mucho camino por recorrer en ese campo.

La mera conexión de ordenadores a una red no garantiza que realicen nada, como tampoco es condición suficiente para una comunicación con sentido entre dos personas la conexión telefónica internacional. Las máquinas deben poder entenderse mutuamente. A nadie sorprenda, pues, que el problema sea tanto más difícil de resolver cuanto mayores sean las diferencias en los soportes lógico y físico de unas y otras.

Las redes de ordenadores no se limitan a comunicar programas y datos al uso. Una entrada de voz puede digitalizarse y almacenarse para una posterior recuperación y entrega. Puede equiparse una red con un sistema de reconocimiento de voz que controle o que facilite la entrada a los programas. En muchos casos, una voz sintetizada puede proporcionar una salida al usuario. En un ordenador facsímil puede incorporarse una entrada junto con otros documentos; y puede transmitirse, almacenarse o procesarse. Las entradas y salidas de vídeo, los gráficos y las imágenes generadas por ordenador pueden, así mismo, formar parte de la interacción mediatizada por una red.

Los orígenes de las redes de ordenadores se remontan al desarrollo de los primeros sistemas de tiempo compartido, al principio de los sesenta, una época en la que un computador era un recurso caro y escaso. La idea que encierra el tiempo compartido es simple. Puesto que muchas tareas, en especial el desarrollo y depuración de programas, requieren sólo una pequeña fracción de la capacidad de un gran ordenador, se sacará mayor rendimiento del ordenador si éste da servicio no sólo a un usuario en un momento determinado, sino a muchos. Para lograrlo, se conmuta entre muchos programas de usuario cada pocos milisegundos. De la gestión de la conmutación se ocupa un programa llamado sistema operativo, que asimismo controla y gestiona las unidades de disco e impresoras y mediatiza toda la asignación de recursos dentro del ordenador.

Del tiempo compartido a las redes hay un escalón intelectual corto. Una vez demostrado que un grupo de usuarios más o menos reducido podía compartir un mismo ordenador, era natural preguntarse si muchas personas muy distantes podrían compartir los recursos —bases de datos e incluso programas especializados— disponibles en sus respectivos ordenadores de tiempo compartido. Con esta idea en la mente, la Oficina de Proyectos de Investigación Avanzada del Departamento de Defensa de los EE.UU. (ARPA, conocida actualmente como DARPA) empezó en 1969 a conectar centros de investigación instalados por todo del país en una red de ordenadores llamada ARPANET. Por la misma época, Tymshare entrevió el valor económico de una red que soportara eficientemente la comunicación interactiva entre terminales remotos y sus ordenadores de tiempo compartido. La red telefónica podía también proporcionar una comunicación entre ordenadores; se empleaba ya para conectar los terminales a los ordenadores remotos. Las compañías telefónicas respondieron a la nueva demanda de redes de ordenadores de gran cobertura y proporcionaron circuitos punto a punto para uso reservado en las redes.

Construida por Bolt, Beranek y Newman, Inc., ARPANET fue la primera red interactiva, ordenador-a-ordena-

dor, de gran cobertura. Como tal, tuvo que vencer difíciles obstáculos técnicos; el más importante: la necesidad de una nueva anchura de banda alta y una estrategia de conmutación de baja demora. La red telefónica emplea la estrategia de conmutación de circuitos: cada llamada se encauza a través de la red en una trayectoria dedicada (reservada o conexión), que se mantiene mientras dura la conversación exactamente igual que si las dos partes en comunicación se hallaran en conexión directa por medio de un circuito único. El inconveniente habitual que lleva aparejado esta solución, y lo comprobará quienquiera que haga llamadas de larga distancia, es que se tardan varios segundos en establecer la conexión. Semejante demora puede resultar irrelevante en el caso de las comunicaciones normales de voz, pero sería intolerable cuando se interpusiera la cascada de rápidos intercambios de datos que suelen producirse entre ordenadores. Una red de rápida conmutación de circuitos podría rebajar esa demora a una fracción de segundo, nivel tolerable para muchos usos.

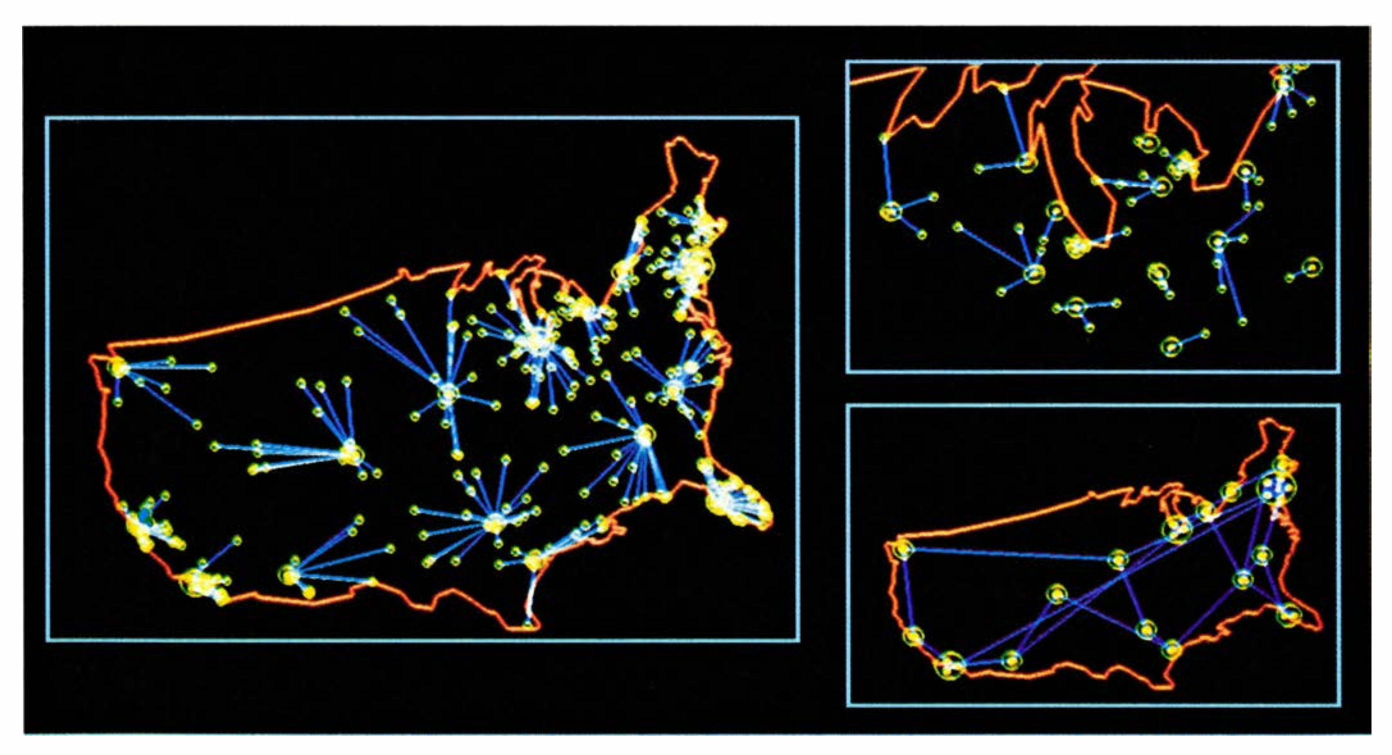
Un circuito reservado puede constituir, sin embargo, un derroche. Imaginemos que aplicamos la conmutación de circuitos a las carreteras y alguien quisiera ir de Barcelona a Zaragoza; en ese orden, habría que avisar a la Dirección General de Tráfico y pedirles que cerraran un carril al resto del tráfico a lo largo de todo el tramo de la autopista A-2.

La solución, en la que ARPANET se convirtió en pionera, fue una técnica llamada conmutación de paquetes. En la conmutación de paquetes la información que se envía de un ordenador a otro no viaja en una corriente continua a través de un circuito reservado, sino que cada ordenador usuario se conecta a un nodo (en el ARPANET original era un minicomputador) y éste reparte los mensajes procedentes del ordenador en paquetes de datos.

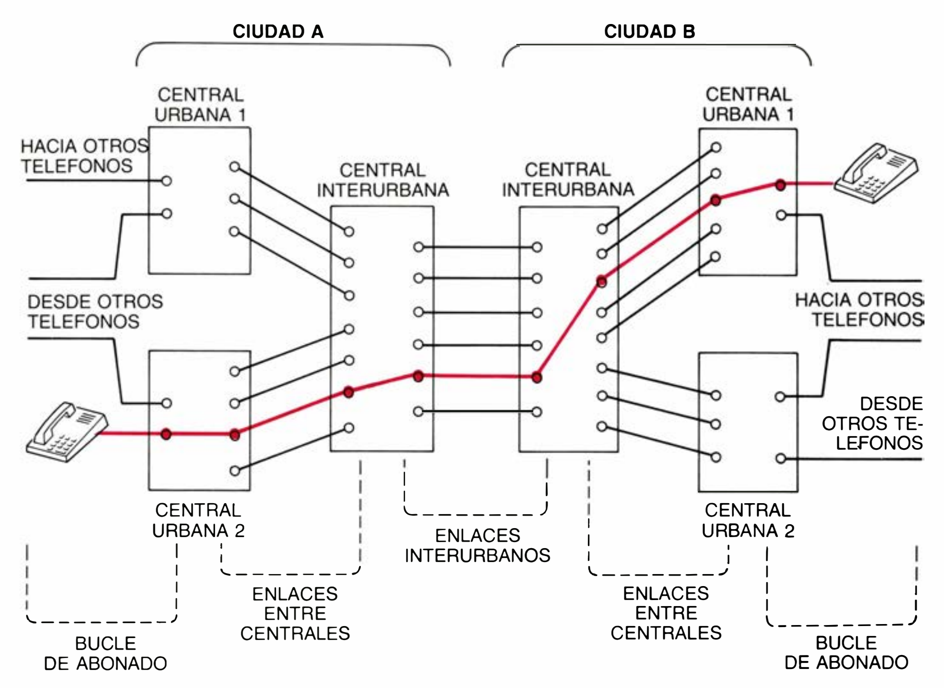
Todo mensaje contiene una secuencia arbitraria de dígitos binarios precedida por alguna información de dirección, y cuya longitud total no supera el tamaño máximo permitido por la red. A cada paquete se le adjunta una "cabecera", que posee, al menos, la dirección de destino y un número de secuencia; los paquetes se envían entonces a través de la red, que actúa como una versión rápida del servicio postal, leyendo las direcciones y entregando los paquetes en una fracción de segundo. No hay demora causada por el establecimiento de circuito, pues no existe tal establecimiento de circuito. En consecuencia, tampoco hay capacidad malgastada, ya que no se produce ninguna conexión individual entre las dos máquinas. Ello no impide que un pequeño porcentaje de la capacidad de comunicación se use para información de rutas o pistas, cabeceras e información de control.

Paquetes sucesivos del mismo mensaje pueden seguir trayectos diferentes para alcanzar un mismo destino. Cuál sea en un momento determinado el trayecto más rápido por la malla lo establecen, conjuntamente, los nodos; y lo hacen de acuerdo con un algoritmo de encadenamiento distribuido. Cada nodo calcula una "métrica de distancia", desde su ubicación hasta los distintos destinos de la red; para ello considera qué líneas están libres, que nodos ocupados, etcétera. Los cálculos se actualizan con cierta frecuencia (por ejemplo, cada medio segundo) y se transmiten a los nodos vecinos. Apoyándose en esta información, un nodo puede cursar un paquete que le llegue hacia el siguiente tramo de su viaje. El algoritmo de ruta del ARPANET original conseguía esto sin conocer la topología ni la condición global de la red. El algoritmo actual distribuye información topológica completa a todos los nodos.

Cuando la comunicación procede en ráfagas cortas y aisladas (por ejemplo, cuando un operador de terminal teclea comandos a un ordenador remoto), la conmutación de paquetes puede sacar-



1. RED DE GRAN COBERTURA, en su fase de diseño; varios estratos del sistema se muestran a la vez en diferentes ventanas. En el mapa de la izquierda se señalan los ordenadores anfitriones y los ensambladores/desensambladores de paquetes (PAD), enlazados en nodos llamados conmutadores de paquetes. El mapa de la región de los Grandes Lagos (*derecha*, *arriba*) muestra terminales particulares enlazados a PAD. En el mapa inferior de la derecha los conmutadores de paquetes se conectan a través de líneas troncales. La imagen se generó mediante el sistema experto DESIGNet, de la BBN Communication Corporation.



2. CONMUTACION DE CIRCUITOS usada por la red telefónica de larga distancia. Para cada llamada, las centrales de conmutación establecen un circuito (*color*) entre dos teléfonos de ciudades distintas. Una parte de la capacidad de transmisión del circuito debe reservarse para el usuario mientras dure la llamada.

le mayor rendimiento a las posibilidades de la red. Al no haber capacidad de transmisión que esté específicamente reservada para un individuo, los intervalos de tiempo muerto entre paquetes pueden llenarse con paquetes de otros usuarios que se dirijan al mismo o a otro destino. De hecho, la conmutación de paquetes representa una clase de sistema multiplexado distribuido, pues permite que todos los usuarios compartan todas las líneas de la red durante todo el tiempo.

Para enlazar un número de nodos (n) en una red de conmutación de paquetes, se necesita un mínimo de $n-1$ líneas; si con una línea adicional se enlazan los dos extremos, la topología de tal malla adquirirá una forma anular. Conviene añadir al anillo conexiones cruzadas. De ese modo, si falla un nodo o circuito, se amortiguará el impacto sobre la red. Puede construirse una red fiable con un número de líneas que sea sólo ligeramente mayor que el de los nodos. De hecho, el fallo de una determinada línea en la red no necesita interrumpir la comunicación entre dos puntos, como ocurriría inevitablemente en una red de circuitos conmutados; los paquetes pueden cursarse por un trayecto alternativo. Es inmediato que debe diseñarse la red para que evite la congestión de tráfico o distribuya rápidamente cualquier sobrecarga.

Ya que los paquetes pueden seguir trayectos diferentes, podrán llegar a su destino fuera de secuencia. Para muchas aplicaciones, esto resulta inaceptable; y así, hay que programar el nodo de destino para que reensamble los paquetes en su orden correcto, antes de aportar el mensaje original. En este caso, el ordenador receptor será incapaz de notar que el mensaje se había fragmentado en paquetes. El ordenador emisor tampoco se entera de la técnica de conmutación subyacente; se limita a proporcionar la dirección de destino al principio del mensaje. Por su parte, el primer nodo conmutador de paquetes adjunta la dirección a cada paquete del mensaje. Al usuario le parece un circuito reservado común: la red suministra un "circuito virtual".

La idea de un circuito virtual hunde sus raíces en el mundo de la telefonía. Sin embargo, en algunas aplicaciones los abonados a una red se sentirían más felices si se movieran en un sistema equivalente al servicio postal. Por ejemplo, cuando hay que transmitir una imagen de un ordenador a otro y se reparten los datos en paquetes, no importa en absoluto en qué orden llegan los paquetes, con tal de que arriben con información suficiente para colaborar en el ensamblado de todos los paquetes y reconstruir la imagen deseada. En ese caso, los "datagramas" de paquetes individuales podrían transmitirse a través de la red como si se tratara de las cartas en el correo; con una salvedad: los datagramas llegarían en contados milisegundos, en vez de los días acostumbrados en el servicio postal.

Desde que ARPANET demostró la viabilidad de la conmutación de paquetes para redes de ordenadores de gran cobertura, han aparecido varios servicios comerciales de redes de paquetes; entre ellos, Telenet de U.S. Sprint, TYNMET de la McDonnell Douglas, también en Norteamérica, P.S.S. en el Reino Unido, etcétera. Todas estas redes constituyen servicios públicos, que operan con líneas reservadas comerciales (las compañías de las redes instalan sus propios conmutadores de paquetes). En su mayor parte, estas redes de gran cobertura ofrecen un servicio de circuito virtual.

El diseño de una red de comunicación dependerá de las distancias a cubrir, el tipo de dispositivos a los que se conectará y las aplicaciones a que se destinará. En el caso de dos o más nodos que se mantengan en proximidad física al medio de transmisión, se tratará, por lo común, de un bus de alta velocidad o de una red de área local. Para redes de gran cobertura, se recurrirá a múltiples enlaces de transmisión y a la de conmutación asociada.

Las redes de área local han conocido últimamente una enorme popularidad. Conectan dos o más ordenadores que, de ordinario, se alojan en un mismo edificio o se agrupan en un asentamiento orgánico —una universidad por ejemplo— para rendir servicio a una misma institución. Además de permitir que los miembros de la institución se envíen mensajes entre sí y accedan a las mismas bases de datos y programas (como es también el caso para las redes de gran cobertura), posibilitan que diferentes estaciones de trabajo compartan un equipo costoso: instalaciones de almacenamiento, impresoras y otras, cuyos precios no han bajado con la celeridad con que lo han hecho los propios de los ordenadores. Se cifra en varios cientos de miles el número de redes de área local tendidas en todo el mundo. Esas redes, basadas en el uso de cable coaxial compartido o de pares trenzados (en vez de complicados conmutadores), suelen proporcionar un servicio de datagramas. Cuando así se precisa, el ordenador de destino realiza la reordenación de los datagramas en secuencia.

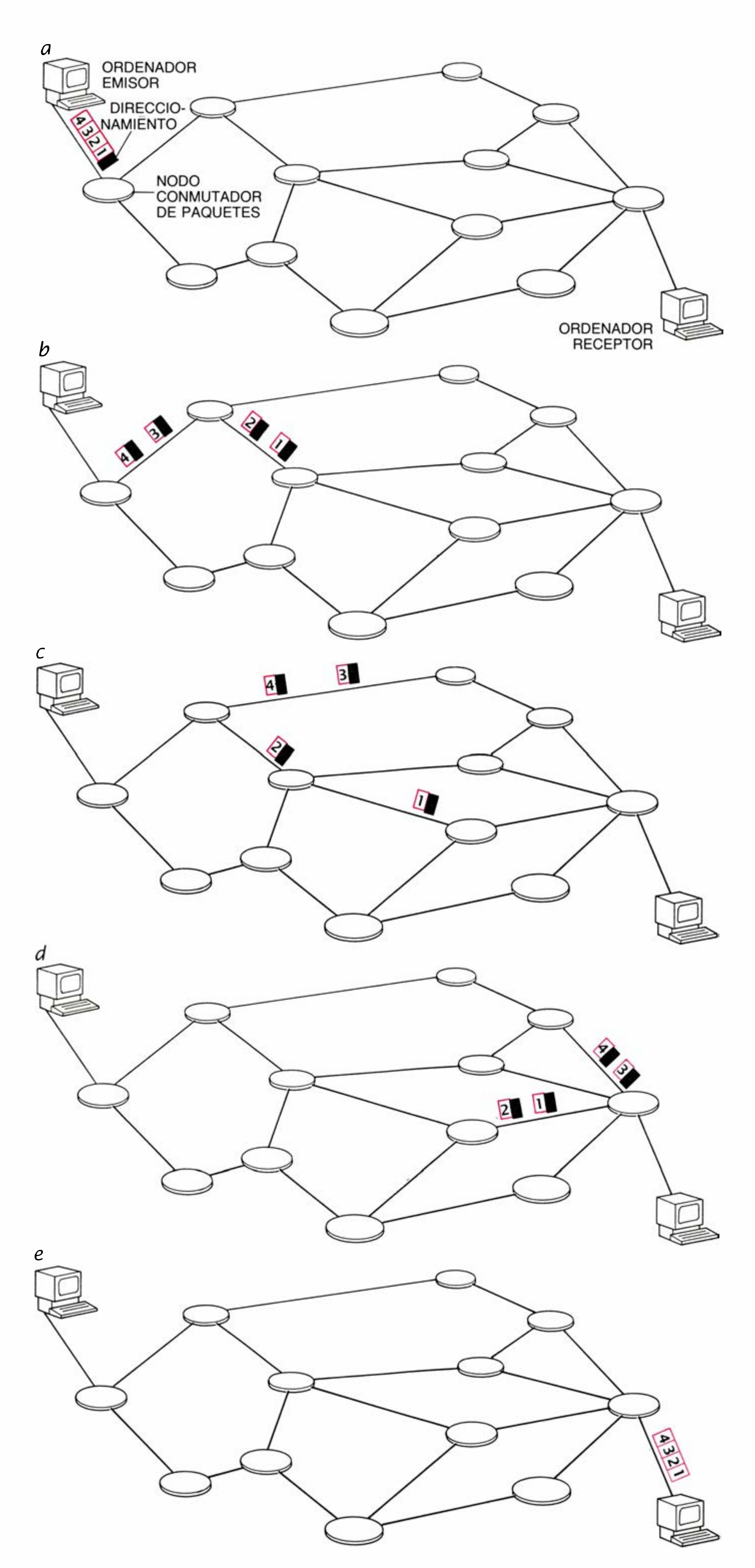
Las dos topologías habituales de redes de área local son el árbol, cuyo prototipo común lo encontramos en la Ethernet de la Xerox Corporation, y el anillo; IBM ha comercializado este último en la forma de anillo con pase de testigo ("token ring"). Ambas son redes de paquetes. El prototipo de Ethernet consiste en un canal único, de

cable coaxial, al que se enlazan los ordenadores, impresoras y otras unidades. Cada máquina se conecta al cable por una interfase que es transmisora y receptora: fragmenta los mensajes de la máquina en paquetes, los envía por el cable y analiza el cable en lo concerniente a los paquetes dirigidos a su máquina.

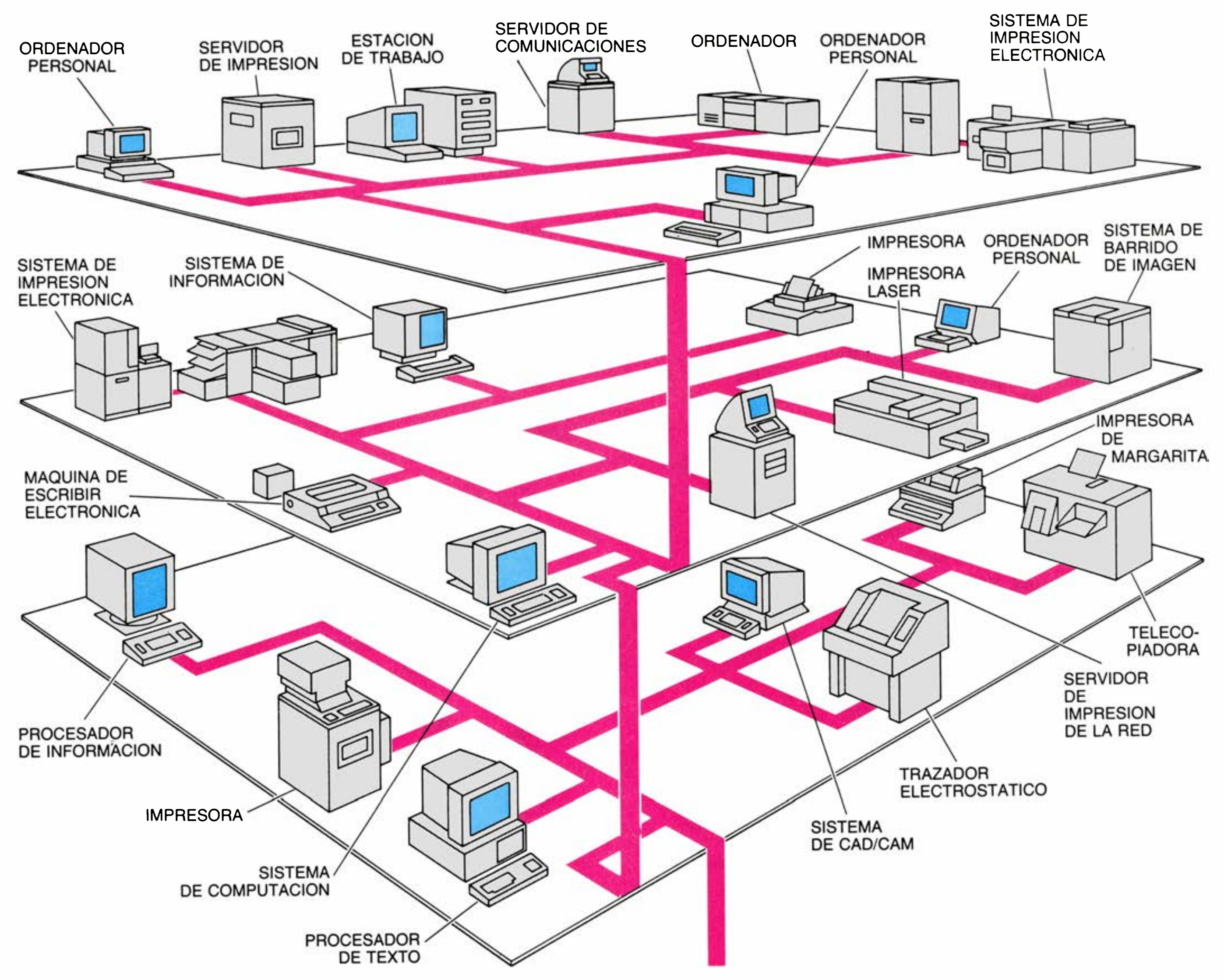
Ethernet es una red de difusión: los mensajes de una máquina se propagan por la red en ambos sentidos a lo largo del cable. De aquí que se pueda presentar un conflicto cuando dos máquinas de la red transmitan al mismo tiempo. En Ethernet, cada interfase tiene un mecanismo que detecta las colisiones de paquetes, aborta la transmisión y la replanifica para momentos después; bajo estas condiciones, pueden darse retardos variables en la comunicación. El sistema funciona bien. No obstante, cuando se trabaja con tasas de transmisión de datos muy elevadas y el número de colisiones crece en demasía puede resultar poco eficiente. Las redes Ethernet actuales transportan unos 10 megabits de datos por segundo.

La red en anillo con pase de testigo es también una red de difusión. Sin embargo, a diferencia de Ethernet, el acceso a la red se organiza por medio de un testigo (una corta secuencia de bits que circula continuamente alrededor del anillo). Para poder transmitir datos, la interfase de una máquina debe esperar primero a que llege el testigo. Aparta luego temporalmente el testigo, inserta en el anillo un paquete con dirección (de una longitud máxima especificada) y reinserta el testigo detrás del paquete. El destinatario propuesto retirará el paquete antes de que el testigo dé una vuelta completa. Así pues, no ocurren colisiones. Ahora bien, según sea la carga de tráfico, la llegada del testigo puede demorarse más de lo usual y producirse retardos variables.

Los anillos con pase de testigo actuales permiten también una tasa de trans-



3. CONMUTACION DE PAQUETES: se desarrolló para las redes de ordenadores. Un ordenador emisor envía un mensaje con dirección (*a*). En el primer nodo de conmutación de paquetes (o en la interfase entre el ordenador y el nodo) el mensaje se divide en paquetes, portador cada uno de una dirección (*b*). Los paquetes viajan independientemente a través de la red. Cada nodo escoge el siguiente nodo de la trayectoria de un paquete, de acuerdo con la información recibida de los nodos vecinos sobre el tráfico, caída de líneas, etcétera. En consecuencia, los paquetes pueden seguir diferentes trayectos (*c*) y llegar descolocados al nodo de destino (*d*). En este punto, el nodo de destino los restablece en su ordenación original antes de la entrega al receptor.



4. UNA RED DE AREA LOCAL permite que las estaciones de trabajo de un edificio o de una zona restringida se comuniquen y compartan recursos. Es el caso de la red Ethernet de la Xerox Corporation, en la cual las máquinas se conectan a través de un cable coaxial o por varios enlazados con repetidores.

misión de 10 megabits de datos por segundo. Se han alcanzado tasas de 100 megabits por segundo o más con anillos experimentales constituidos por fibra óptica, a cuyo través se transporta la información por pulsos de luz en vez de las acostumbradas señales electrónicas. La tecnología óptica podría incluso superar el alcance de las redes locales.

La comunicación entre tipos idénticos de ordenadores presenta ciertas ventajas naturales, pues el soporte lógico del sistema que gobierna ambas máquinas puede ser el mismo. No es fácil conseguir que tipos distintos de máquinas que trabajen juntas ajusten sus diferencias de soporte físico para operar con una programación común. Aunque se ha progresado mucho en el desarrollo de técnicas para intercambiar datos entre distintos modelos de máquinas, persisten todavía no pocas barreras. Más difícil resulta todavía poder intercambiar soporte lógico entre sistemas heterogéneos.

Los programas que se intercambian entre tipos idénticos de máquinas parten con mucha mayor probabilidad de poderse ejecutar, ya que las instrucciones subyacentes son las mismas y se usan los mismos sistemas operativos de soporte lógico. Pero cabría que un programa pudiera desarrollarse sin compilación ni modificación de ningún tipo en un punto remoto y no haber memoria suficiente en el lugar de destino para prestarle el debido soporte. Los requisitos de compilación de un programa en un lenguaje de alto nivel podrían quizá satisfacerse adecuadamente en una máquina, pero no en otra. Sólo cuando las máquinas tienen idéntico soporte lógico resulta verosímil el éxito en la transferencia y ejecución de un programa.

Las funciones reticulares más potentes se han venido limitando a sistemas de máquinas homogéneas que gozan de una administración común. Ejemplos de tales funciones son los sistemas de ficheros distribuidos (que facilitan a una estación de trabajo el acceso inmediato a ficheros, con independencia de su ubicación en la red) o los programas en migración (que viajan por la red hasta el lugar donde han de ejecutarse). Ello contrasta con otra realidad: muchas de las interacciones conseguidas entre máquinas heterogéneas son todavía bastante elementales (envío de mensajes o ficheros de una máquina a otra, ejecución de un programa lanzado de un terminal remoto y otros por el estilo). Ocurre que la heterogeneidad de los ordenadores es la situación dominante, y ni siquiera es raro que una misma institución compre equipos de diferentes fabricantes. De ahí la premura en vencer las barreras que impiden la comunicación entre máquinas heterogéneas.

Consideremos un ejemplo sencillo de una barrera de ésas. El ordenador del ejemplo almacena datos en palabras de 32 bits; en otro, la longitud de la palabra es de 16 bits. Para comunicarse las dos máquinas, el ordenador receptor deberá conocer cómo disponer los datos en su memoria. Ya que la longitud de la palabra en ambos casos es un múltiplo entero de caracteres de 8 bits (octetos o "bytes"), la tarea resulta bastante fácil. Se torna más complicada cuando un ordenador de 32 bits ha de comunicarse con otro de 36 bits: en ese caso, habrá que segmentar algunos octetos o dejar espacios vacíos en la memoria. Más aún: los datos no pueden transmitirse aislados, sin una suerte de "carta de presentación" que los describa o los interprete.

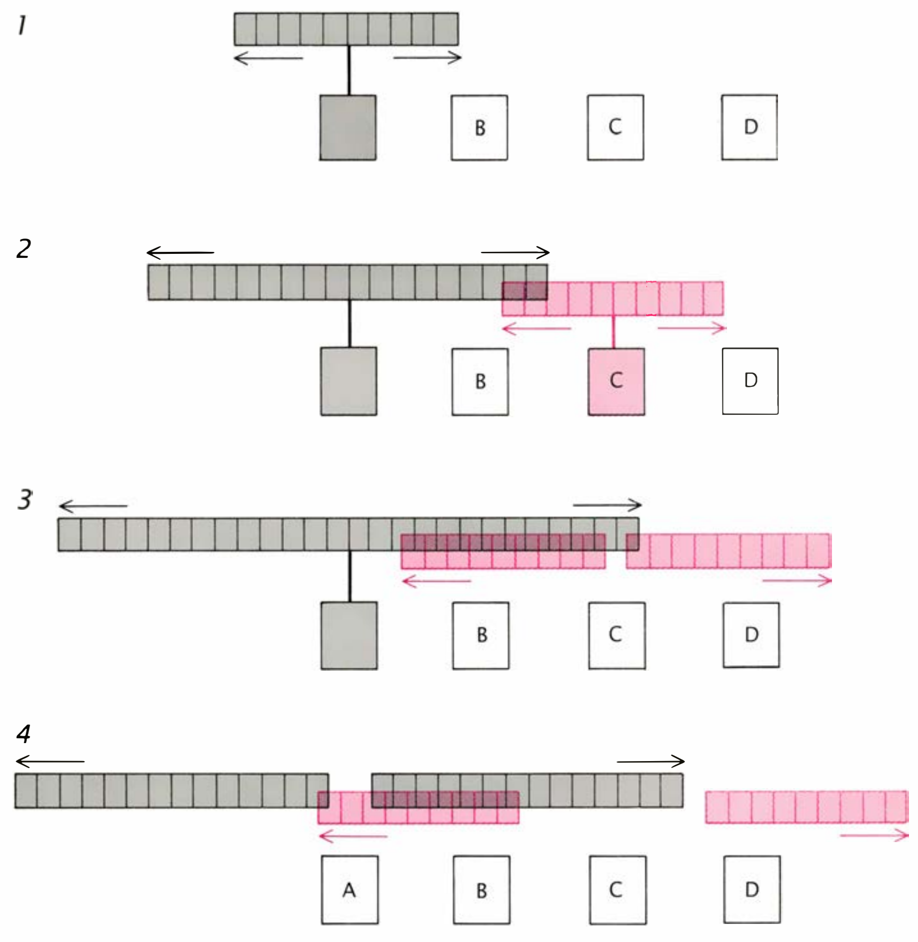
Cuando lo que se ha de transmitir es un programa, y no datos inejecutables, puede resultar incluso mayor la necesidad de aportar información de contexto que posibilite la operación del programa en un sitio remoto. Por "contexto" podría entenderse el entorno lógico general donde el programa ha de funcionar, incluyendo de manera preeminente el sistema operativo. En principio, podría transmitirse, entero, el soporte lógico junto con el programa; en la práctica, esto resulta, sin embargo, difícil de alcanzar: muchos sistemas operativos están diseñados para desenvolverse en un determinado sorporte físico y no pueden instalarse en otros modelos de máquinas. Sea, por ejemplo, el sistema operativo UNIX. Escrito en un lenguaje de alto nivel, el sistema UNIX puede transitar entre algunas máquinas merced a la fácil obtención de compiladores adecuados. No obstante, se necesita cierta instalación o hechura a medida para adaptar el UNIX a las características de una máquina específica. Una solución sería crear una "arquitectura de máquina virtual" que permitiera que las máquinas heterogéneas limaran las diferencias de su estructura o soporte físico. Una arquitectura de microcódigo normalizada posibilitaría la exploración de varias arquitecturas de máquina virtual.

Una agrupación de ordenadores que trabajen juntos forman un sistema distribuido. El intercambio de información de contexto entre estos ordenadores implica una serie de protocolos que deben ser estrictamente observados por las diversas máquinas participantes para señalar la clase de información de contexto y cómo interpretarla. Estos protocolos son a los sistemas distribuidos lo que los lenguajes de programación son al soporte lógico. Permiten especificar la arquitectura global y la operación de los elementos críticos del sistema.

Un planteamiento prometedor en el que se está trabajando internacionalmente es el de la creación de un entorno común de comunicaciones para sistemas distribuidos; se pretende de este modo que los vendedores de soporte físico y lógico, así como los diseñadores de redes, se adhieran a los protocolos. Los protocolos establecen una forma de arquitectura de comunicación "abierta" de la red.

Desde hace algún tiempo, la Organización Internacional de Pesos y Medidas cuyas siglas, ISO, corresponden a su denominación inglesa, se viene ocupando de la delimitación de protocolos de red que den base a la Interconexión de Sistemas Abiertos (OSI). Para la mejor resolución del problema se ideó un modelo conceptual que divide las actividades de una red en siete capas o niveles, a cada una de las cuales se asocia un protocolo o protocolos específicos. Por ejemplo, la capa inferior es el nivel físico que contiene las normas para las convenciones de señalización, las velocidades de transmisión, etcétera. Cerca de la superficie del modelo, el nivel de "presentación" estipula cómo han de interpretarse los datos para su presentación; un ordenador con el programa de traducción adecuado estará dotado para conversar con otros ordenadores que conforman sus datos de modo distinto. Por último, los protocolos de la capa superficial especifican los tipos de mensaje que pueden enviarse a través de la red para llevar a cabo las diversas aplicaciones.

Las interacciones que la actual arquitectura OSI posibilita se limitan principalmente a la transferencia de datos. No dicta cómo los ordenadores de la red deberían colaborar en aplicaciones

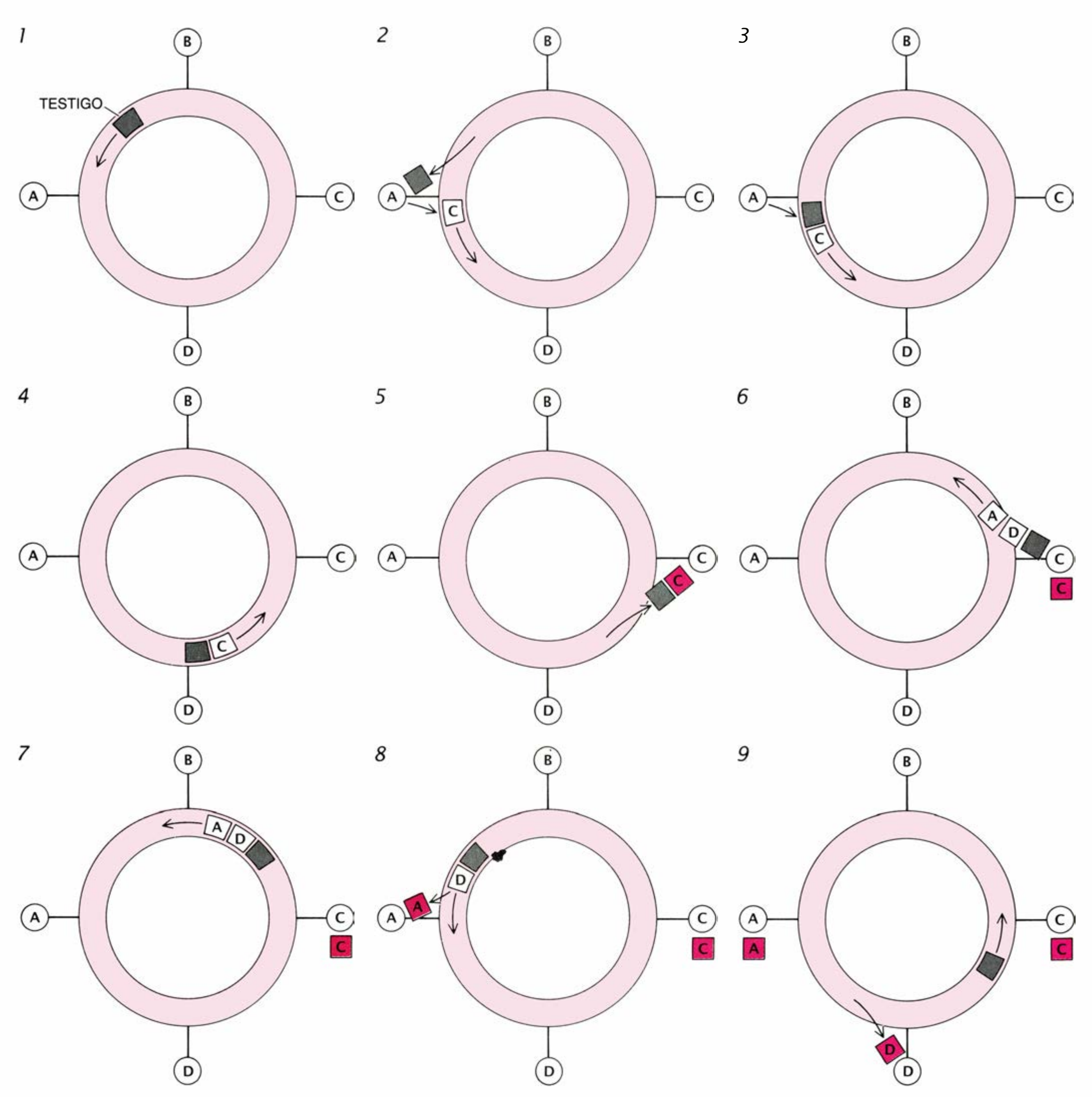


5. DETECCION DE COLISIONES. Se trata de una característica esencial de la red Ethernet. Puesto que los paquetes viajan en ambos sentidos de la dirección de un cable coaxial, dos estaciones pueden empezar a transmitir simultáneamente. Para evitar la confusión, cada estación de Ethernet posee un mecanismo de detección de colisiones. En la secuencia que se muestra aquí, la estación *A* empieza a transmitir un paquete a la estación *D* en un momento en que la red estaba en reposo (*1*). Sin embargo, antes de que el paquete llegue a *C*, éste empieza a enviar su propio paquete (*2*). Los dos paquetes colisionan, y cuando *C* detecta el choque, interrumpe la transmisión (*3*). Poco después, el paquete de *C* llega a *A*, que detendrá la transmisión (*4*). Ambas estaciones de la red *Ethernet* esperan un corto intervalo aleatorio y vuelven a intentarlo de nuevo.

de alto nivel. Mas, a pesar de sus ambiciosos objetivos, cuesta llevar a cabo la estrategia de la arquitectura abierta. A grandes rasgos, el problema estriba en que las especificaciones de un protocolo son inevitablemente ambiguas y cada vendedor las interpreta a su modo. Una simple diferencia en la interpretación puede ser (y así ha sido) suficiente para hacer incompatibles los productos resultantes.

En 1985 un grupo de casas comerciales fundó la Corporación para Sistemas Abiertos (COS), con el propósito de ayudar a resolver las incompatibilidades en la realización del protocolo OSI. Dicha corporación equivale a un laboratorio de certificación: su cometido consiste en analizar los productos y darles un marchamo de validación si son compatibles con otros productos certificados. Los ensayos no son, por supuesto, del todo seguros. Dos productos que sean compatibles con la batería de pruebas del COS pueden seguir sin lograr trabajar juntos correctamente. Por eso, antes de que un producto reciba el certificado de un grado de confianza razonable, ha de examinarse con varios sistemas certificados más. En contraste, la amplia popularidad y uso del protocolo de transporte TCP/IP en EE.UU. (que fue anterior al protocolo de transporte OSI) se debe a su probada eficacia en las instalaciones donde se ha ensayado.

Mientras tanto, las redes proliferan y la cacofonía aumenta. El problema de la compatibilidad vuelve a presentarse



6. ANILLO CON PASE DE TESTIGO ("token ring"). Es éste otro diseño habitual de redes locales. El acceso a la red se controla mediante una secuencia circulante de bits llamada testigo. Para enviar un mensaje, una estación ha de esperar que le llegue el testigo, apartarlo del anillo, poner un paquete con la dirección debida (o una secuencia de paquetes) en el anillo y colocar el testigo detrás del paquete en cuestión. Una estación puede retirar un mensaje dirigido a ella sin tener que separar el testigo. En la ilustración *A* envía un mensaje a *C*, que lo recibe y se encarga, a su vez, de enviar sus propios mensajes hacia *A* y *D*.

en otro nivel: no sólo las máquinas de una red determinada deben poder entenderse entre sí; han de poder hacerlo las propias redes. Las empresas y otras instituciones necesitan conectar sus redes locales a redes de gran cobertura o enlazar varias redes locales juntas. Pero puede ocurrir que tales redes difieran en los formatos de datos, las longitudes de paquetes, las velocidades de transmisión (las redes locales suelen ser mucho más rápidas que las de gran cobertura), los mecanismos de dirección y así sucesivamente.

El protocolo de transporte anfitrión de ARPANET se diseñó, en su versión original, con la idea de que funcionase como si se tratara de un buen periférico de ordenador. Un supuesto que hoy sobra, pues hay ya muchas clases de redes, como las redes por radio, que no siempre logran hacer llegar todos los paquetes.

Los investigadores de DARPA adoptaron la solución de desarrollar una arquitectura inter-reticular, que gozaría de una amplia difusión en las redes de área local y serviría de modelo en múltiples situaciones. El protocolo de interfase de la red CCITT X-25 especifica una vía alternativa para la interconexión reticular que presupone el establecimiento de un circuito virtual entre redes. Suele usarse para conectar redes (circuitos virtuales) dotadas de una velocidad similar de transmisión y no requieren, por tanto, ninguna puerta especial ("gateway").

El diseño inter-reticular incorpora el protocolo de transporte TCP ya mencionado y puertas para enlazar redes. La puerta constituye un nodo especial creado para encapsular el texto de los mensajes que llegan de una red en la forma que demanda la otra, y viceversa. Se sirve de un protocolo entre-redes (IP) para interpretar las direcciones de ámbito inter-reticular, para segmentar los paquetes cuando sea necesario o facilitar, si de eso se trata, el transporte de los paquetes entre redes muy dispares. Por ejemplo, si un paquete de 500 bytes llega por una puerta a una red cuya longitud máxima de paquete es de 250 bytes, la puerta parte el paquete en dos, les fija la dirección a ambas mitades y las encamina (quizás independientemente) hacia su destino. Las puertas pueden comunicar, asimismo, con otras puertas y formar un sistema inter-reticular que reparte los mensajes a través de redes múltiples. Actualmente OSI incorpora el IP como mecanismo inter-reticular.

En el desarrollo de la arquitectura inter-reticular hubo un propósito inamovible: que la participación de cualquier red no requiriera ningún cambio interno. Ello implicaba dotar a las puertas de las siguientes propiedades: capacidad para pasar paquetes de unas redes a otras con distinta conformación externa e interna (por ejemplo, diferente velocidad de trasmisión, tamaño máximo de paquetes, formatos y mensajes de error) y capacidad para comunicarse con otras puertas durante la circulación inter-reticular de los mensajes.

Un área de importancia prometedora en el futuro de las redes es la relativa a la comunicación y computación portátil. Cierto es que la mayor parte de las necesidades se ubican hoy en el entorno de la empresa o del hogar; pero no lo es menos que se abre un enorme potencial de transformación de la sociedad si extendemos ese tipo de comunicaciones al entorno móvil. Contamos ya con las bases para iniciar un proceso de cálculo en un lugar, continuarlo en el avión con un ordenador portátil de regazo y completarlo en otro punto geográfico. Tales son los primeros pasos: comunicaciones móviles a través del teléfono celular, redes de paquetes vía radio, dispositivos de bandas civiles y de radioaficionado, así como los de página. La propagación multitrayecto y otras limitaciones técnicas impedirán que la tasa máxima de transferencia de datos en un entorno móvil exceda de unos pocos megabits por segundo. Cifra que es suficiente, pues muchas aplicaciones importantes de los ordenadores y las comunicaciones requieren sólo una moderada capacidad de proceso y una comunicación asimismo restringida. Además, se están explorando nuevas aplicaciones para servicios que precisen una velocidad superior.

La mayor capacidad de las redes ha venido instada, casi siempre, por la necesidad de prestar servicio a nuevos usuarios. A pesar de que la demanda de capacidad de comunicación de cada usuario se mantenga más o menos constante, la suma de las demandas de todos ellos crece. Y eso fue lo que estimuló la evolución del sistema telefónico analógico. Los canales individuales de voz se agregaron en grupos de canales de voz, y estos grupos se unieron, a su vez, en supergrupos de más alto nivel. De un modo parecido, las demandas de transmisión digital de un mismo usuario pueden agruparse, ya sea por la empresa que ofrece el ser-

vicio o por el propio usuario. Tal es el fundamento de la Red Digital de Servicios Integrados (ISDN) que acaba de constituirse; en ella el usuario decidirá cómo multiplicar los datos digitales en uno o varios circuitos. Esa canalización múltiple es sólo una dimensión del tema de la capacidad.

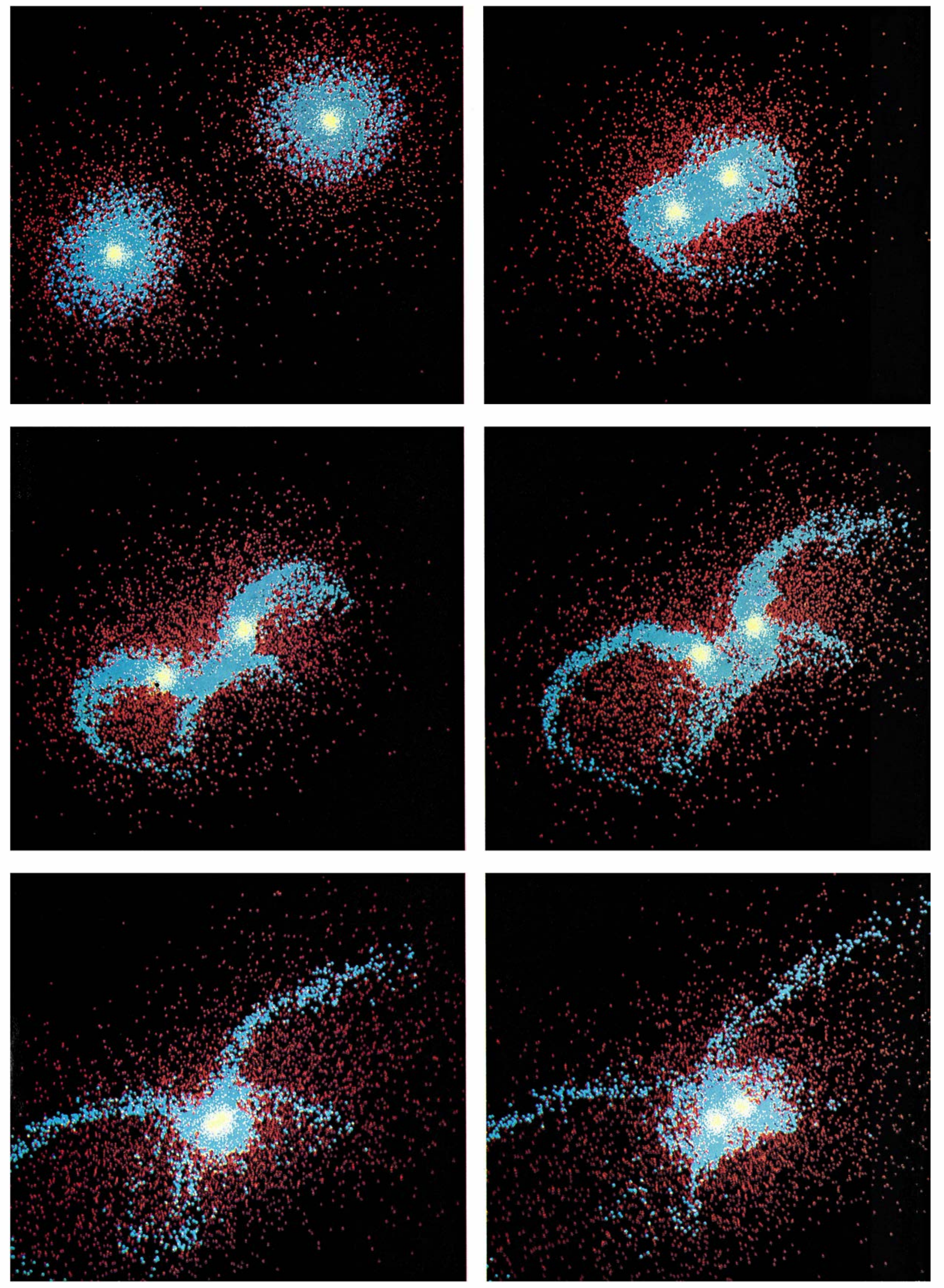
En el futuro, habrá cambios sustanciales en la oferta de tipos de servicios de las redes y en sus costes. En parte, los cambios vendrán determinados por la expansión de la capacidad transmisora de las líneas de comunicación. Las tarifas actuales se basan principalmente en el coste de los canales individuales de voz. Si una línea puede soportar 24 canales de voz, y el grueso del negocio de la compañía telefónica se centra en la venta de canales de voz, es de suponer que el coste se abaratará. Los circuitos ópticos, múltiples, resultarán prohibitivos para muchos usuarios, a no ser que se produzca una drástica reclasificación de las tarifas. Reclasificación que sólo será posible si el sector de los sistemas de comunicaciones evoluciona hacia dispositivos que requieran más alta velocidad, el vídeo por ejemplo. Si las compañías telefónicas llegan a vender tantos circuitos de vídeo como canales de voz, al mismo precio más o menos, no perderán dinero y se podrán establecer las bases de unos circuitos de muy alta velocidad. Pero no se trata de una empresa nimia, y el éxito dependerá de que la sociedad genere una demanda importante.

Cuando se construyó la primera ARPANET en 1969, se apoyaba en líneas reservadas de alta velocidad que transportaban 50 kilobits por segundo. La Fundación Nacional estadounidense para la Ciencia (NSF) ha establecido una red de superordenadores, NSFNET, que tiene parecidas tasas de transferencia de datos. La NSF prevé arrendar líneas de más alta velocidad en un futuro próximo para satisfacer las necesidades a largo plazo de la comunidad científica de uso de redes. Se dispone ya de líneas digitales que operan a una velocidad de 1,5 megabits por segundo, y se ha introducido recientemente un servicio de 45 megabits por segundo. Una vez completado el tendido de fibras ópticas que se halla en curso en los EE.UU., tales cifras parecerán pequeñas. Se han alcanzado en pruebas de campo tasas de transmisión por encima de un gigabit por segundo; hacia finales de siglo, podrían lograrse velocidades de 10 gigabits por segundo. La canalización múltiple de señales de diferente longitud de onda posibilitaría tasas de transmisión por encima de los 100 gigabits por segundo sobre una fibra única, y hacerlas rentables.

No es difícil pensar qué cosas lograríamos con semejante capacidad. Podría comunicarse, muy rápidamente, una gran base de datos con otros ordenadores; las simulaciones videodinámicas, habituales en la investigación científica, podrían transmitirse desde el superordenador hasta estaciones de trabajo remotas. Imágenes, voz y datos podrían enviarse simultáneamente de una oficina a otra de una misma empresa o institución.

¿Cuánta capacidad de transmisión requerirían tales actividades? En un terminal de ordenador con una pantalla cuadrada de 1000 pixeles por lado, cada píxel de una imagen en blanco y negro se codifica con un bit. La transmisión de 30 imágenes por segundo (para evitar el parpadeo) arrojaría un flujo de información de 30 megabits por segundo. Las imágenes en color requerirían 16 bits de información por píxel; se necesitarían, pues, 480 megabits por segundo para la transmisión en tiempo real de una imagen en color y movimiento. Cuando obtengamos dispositivos de representación visual con pantallas mayores (de 1,2 por 1,8 metros), las exigencias de transmisión aumentarán en un factor de 10 o incluso de 100 si se precisa una operación en tiempo real. Podrán usarse, obviamente, técnicas de compresión de datos para amortiguar tales exigencias. Pero debería quedar claro que las redes de ordenadores tienen muchas aplicaciones potenciales en el dominio de la transferencia de información ofrecida por las fibras ópticas.

La cuestión se torna más evidente en cuanto nos detenemos en la función de los ordenadores. No son meros portadores de la información: pueden procesarla e incluso "razonar" sobre ella. Razonar sobre una imagen requerirá, a buen seguro, bastante más información que la que está contenida en la propia imagen, información contextual que deberá transmitirse con la imagen. Lo mismo cabe decir a propósito de la información que no aparece en forma de imagen: un documento o un diseño VLSI; para razonar sobre ellos, el ordenador debe contar con un notable conocimiento contextual. Las fibras ópticas incorporadas en las redes de ordenadores posibilitarán, ésa es la esperanza, los niveles de flujo de información necesarios para que las redes cumplan su próposito final, esto es, permitir a los ordenadores colaborar inteligentemente en la solución de los problemas humanos.



Computación avanzada en el campo de la ciencia

Los experimentos informáticos están enriqueciendo la investigación científica. Su importancia resulta hoy equiparable a la que revisten las teorías, las observaciones y los experimentos de laboratorio

Piet Hut y Gerald Jay Sussman

La computación de alta velocidad está cambiando de modo espectacular la forma de "hacer ciencia". Tradicionalmente, los investigadores vienen desarrollando modelos hipotéticos a partir de los cuales se avanzan predicciones contrastables mediante la observación o la experimentación. En el caso de sistemas complejos, resulta difícil encontrar simplificaciones que sigan representando adecuadamente la realidad, pues algunos de los fenómenos más importantes derivan de la combinación o multiplicación imprevisible de pequeños efectos.

El empleo de ordenadores permite ahora a los investigadores abordar con éxito fenómenos que abarcan desde la difusión de portadores de carga en un semiconductor hasta las colisiones entre galaxias preñadas de millones de estrellas. El desarrollo de modelos de esos sistemas posibilita poner sobre el tapete qué consecuencias derivan de la aplicación de teorías simples a situaciones complejas. Además, si se dispone de suficiente potencia de cálculo, cabe determinar las consecuencias de una teoría sin recurrir a aproximaciones o simplificaciones cuestionables. Un modelo informático suficientemente preciso permite la realización de mediciones comparables a las habituales en el laboratorio, e incluso resulta preferible a éstas, pues con él se llevan a cabo ensayos impracticables y se miden parámetros inaccesibles a la experimentación o a la observación. Demos un ejemplo; gracias a nuestro conocimiento preciso de la ley de Newton de la gravitación de los cuerpos y los modelos informáticos ajustados del sistema solar, podemos determinar cómo variaría la órbita de la Tierra si Marte no existiese.

Los teóricos se han visto a menudo obligados a simplificar en exceso, para calcular efectos aproximados y compararlos con el mundo real. Una de las consecuencias del incremento del poder de cálculo puesto a disposición de los científicos ha sido el que los métodos de "análisis por síntesis" han ganado terreno frente a los métodos reduccionistas. Se recurre al enfoque sintético cuando se conocen los procesos fundamentales de las interacciones entre las partes de un sistema, pero no la configuración detallada de éste. Si se trata de determinar la configuración desconocida por síntesis, podemos examinar las configuraciones viables y averiguar las consecuencias de cada una de ellas. Haciendo corresponder cuidadosamente los datos registrables de la situación experimental con esas consecuencias, se escoge la configuración que mejor explique nuestras observaciones.

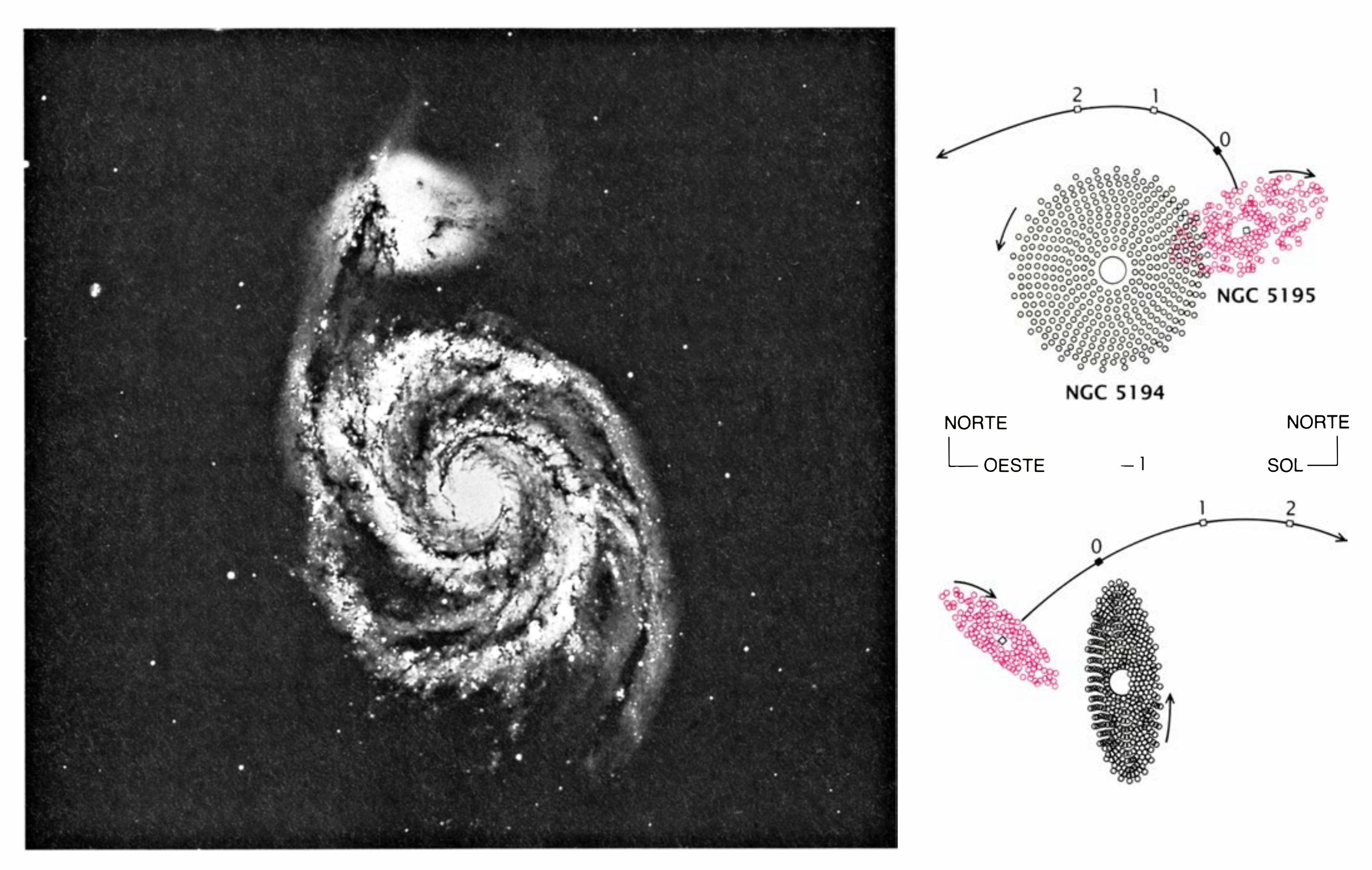
Un famoso ejemplo del enfoque sintético es el de los esfuerzos realizados, en el siglo XIX, por entender las perturbaciones de la órbita de Urano; nadie acertaba a justificarlas. Los investigadores añadieron un planeta hipotético al sistema solar y variaron los parámetros de su órbita hasta que se obtuvo una reconstrucción satisfactoria de la perturbación. El ensayo llevó directamente al descubrimiento de Neptuno, encontrado cerca de la posición predicha. En el pasado, el enfoque sintético se limitaba a situaciones bastante sencillas. Merced a los instrumentos de cálculo de alta velocidad, los métodos sintéticos ocupan hoy un lugar de privilegio al lado de los tradicionales métodos reduccionistas de análisis.

El análisis por síntesis ha arrojado nueva luz en uno de los campos tradicionales de la astrofísica, la dinámica gravitatoria; el incremento de la velocidad de los ordenadores ha permitido a los investigadores atacar problemas que durante mucho tiempo resultaron inabordables desde otros enfoques.

Considerado a gran escala, el universo está poblado por enjambres de galaxias, algunos de los cuales flotan más o menos solitarios, mientras que otros lo hacen en cúmulos y en supercúmulos. Cada galaxia consta de un número ingente de estrellas y grandes cantidades de gas y polvo. Las distancias entre las estrellas de una misma galaxia son, respecto de sus diámetros, lo suficientemente grandes para que las colisiones mutuas resulten rarísimas, a pesar de sus velocidades, cifradas en decenas o centenares de kilómetros por segundo. En una escala mayor, las colisiones entre galaxias abundan más, pues las distancias intergalácticas son sólo 10 o 100 veces mayores que sus diámetros. En escalas aún mayores, los cúmulos están separados entre sí por distancias equiparables a sus propios diámetros; las interacciones entre ellos se produjeron durante períodos comparables a la edad actual del universo, por lo que no pueden estudiarse como sucesos aislados.

Las colisiones entre galaxias constituyen uno de los "accidentes de tráfico" más espectaculares del universo.

1. SECUENCIA DE IMAGENES GENERADAS POR ORDENADOR que simulan una colisión entre dos galaxias espirales; la figura superior derecha se parece a la pareja de galaxias denominadas de las antenas (o NGC 4038 y NGC 4039) de la constelación del Cuervo. Los puntos amarillos representan los núcleos de las galaxias y, los azules, los discos de estrellas que los rodean. Los puntos rojos simbolizan el halo de materia oculta de las galaxias, que no se ve pero que se manifiesta por sus efectos gravitatorios. La simulación, que abarca un período de cerca de 500 millones de años, es obra de Joshua Barnes, del Instituto de Princeton.



2. CHOQUE ENTRE GALAXIAS, simulado en el experimento, hoy clásico, que realizaron en un ordenador, a principios de la década de 1970, Alar Toomre, del Instituto de Tecnología de Massachusetts, y Juri Toomre, a la sazón en la Universidad de Nueva York. La simulación desarrolla el choque de la "nebulosa torbellino" (también conocida por galaxia NGC 5194, o M51) y la galaxia NGC 5195 (*izquierda*). Los Toomre escogieron varias condiciones iniciales después de un cuidadoso análisis de datos obtenidos por observación. Para cada conjunto de condiciones, desarrollaron una secuencia de "instantáneas". Una de las se-

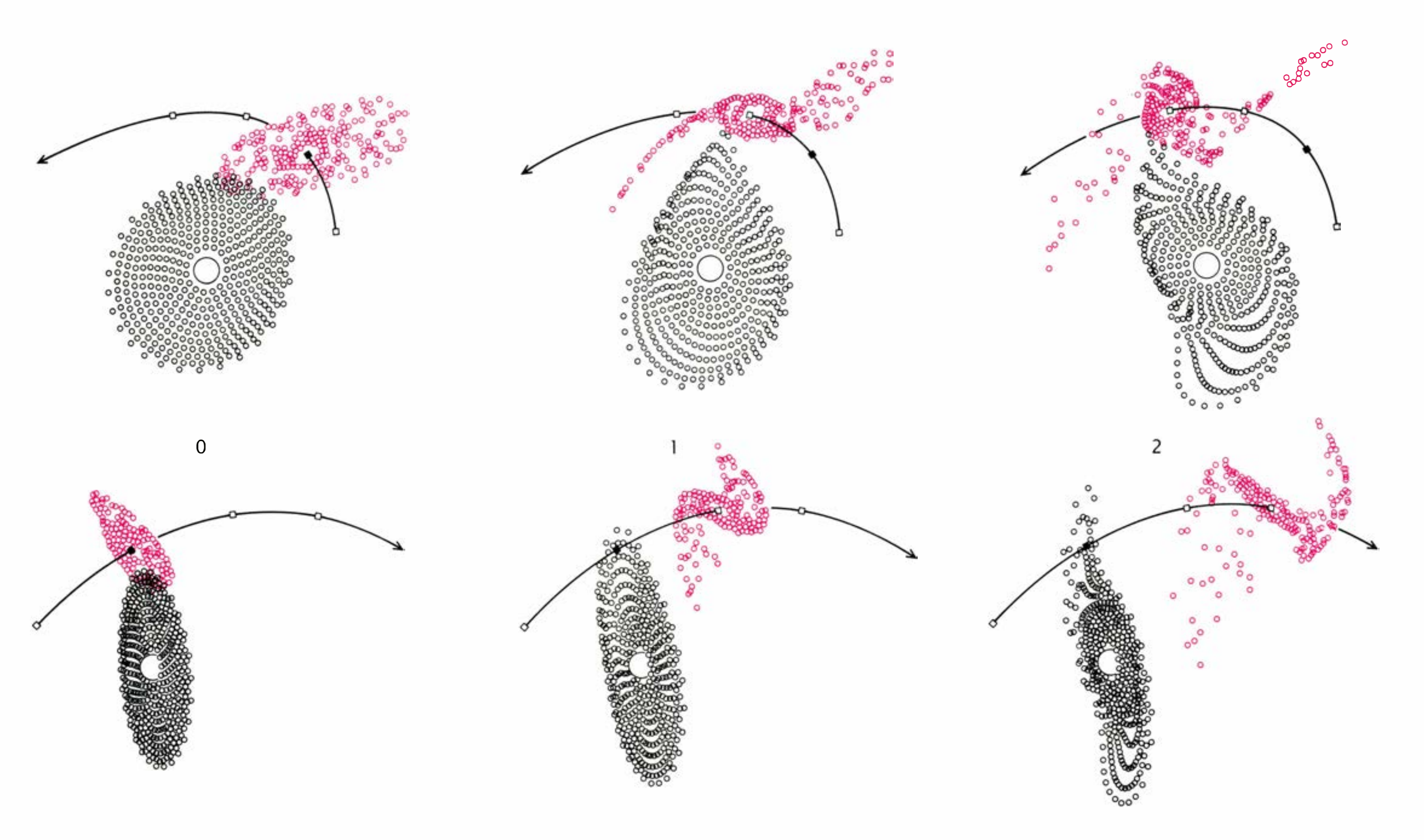
La imagen del cielo que proporcionan los telescopios brinda abundante documentación de tales encuentros. Se conocen galaxias apareadas muy próximas entre sí y que parecen estar unidas por puentes. Mientras que la mayoría de las galaxias presentan un diseño aproximadamente espiral o una forma esférica o elíptica sencilla, las galaxias que parecen estar implicadas en colisiones muestran deformaciones o irregularidades. En la década de 1950 se sugirió que las deformaciones que sufren las galaxias que interactúan debíanse a las fuerzas gravitatorias que ejercen unas sobre otras. En la de 1960 gozaron de popularidad hipótesis más exóticas, en las que se hacía intervenir campos magnéticos y explosiones. El rechazo de las explicaciones gravitatorias, más sencillas, respondía a la dificultad de entender que las interacciones meramente gravitatorias generaran estructuras tan prominentes como las complejas colas y puentes que emanan de los pares de galaxias en interacción, y que engarzan a veces unas con otras.

A principios de los años 70, Alar Toomre, del Instituto de Tecnología de Massachusetts, y Juri Toomre, entonces en la Universidad de Nueva York, llevaron a cabo la primera modelización sistemática de los choques entre galaxias. Sus resultados fueron sorprendentes: mediante el empleo exclusivo de fuerzas gravitatorias, reprodujeron las formas observadas en varios sistemas de galaxias en interacción. Los Toomre equipararon el campo gravitatorio de una galaxia al campo asociado a una masa puntual situada en el centro de la galaxia. Simularon los discos de las galaxias, formados por decenas de miles de millones de estrellas, mediante unos escasos centenares de partículas que se movían en el campo de las galaxias del modelo, sin intervenir en dicho campo. Incluso con esas burdas aproximaciones, los investigadores comprobaron que los encuentros entre las galaxias del modelo generaban una deslumbrante variedad de patrones, a modo de "salpicaduras" de la colisión entre dos masas fluidas, que a menudo formaban láminas y cintas. Contempladas de perfil, las láminas y cintas reproducían algunos de los rasgos, delicados y bien definidos, que habían instado a otros astrónomos a invocar complicados procesos magnéticos y de confinamiento.

El trabajo de los hermanos Toomre brinda un claro ejemplo de cómo se obtiene nueva información mediante síntesis. La secuencia de "instantáneas" de ordenador de la figura 2, por ejemplo, muestra que la flámula que aparentemente conecta la "nebulosa torbellino" (también denominada galaxia NGC 5194, o M51) con la galaxia NGC 5195 no constituye ningún puente. La conexión es mera ilusión, provocada por la perspectiva con que se contemplan las galaxias en el firmamento.

Muchos problemas de física, demasiado complejos, se resisten a la síntesis directa en modelos globalizadores. Pero cabría quizás una aproximación en dos etapas, que parta de la simulación de secciones microscópicas de un sistema grande. Los métodos sintéticos ofrecen a menudo inferencias cuantitativas sobre el comportamiento estadístico de subcomponentes muy pequeños de un sistema macroscópico. Podría recurrirse entonces a un modelo de nivel superior, que combinase esas reglas estadísticas, cual piezas básicas, para predecir el comportamiento del sistema entero.

Ejemplo de problema que excede la capacidad de cálculo actual es la mo-



cuencias produjo una instantánea (*derecha*) que semeja notablemente la configuración real de las galaxias. Las dos filas representan dos perspectivas distintas del mismo acontecimiento; la separación entre instantáneas es de 100 millones de años. El experimento de los Toomre demuestra que la flámula que aparentemente conecta M51 y NGC 5195 no constituye ningún puente. La conexión aparente es, en realidad, mera ilusión, provocada por la perspectiva con que se contemplan las galaxias en el firmamento. En estos momentos, la galaxia NGC 5195 se encuentra completamente detrás de M51 y la colisión es ya historia.

delización de la evolución de un cúmulo estelar formado por un millón de estrellas: ni más ni menos que el tamaño de algunos cúmulos globulares en órbita alrededor de nuestra galaxia. Se ha comprobado que puede obtenerse información sobre tales cúmulos mediante un enfoque microscópico que estudia las interacciones entre una estrella solitaria y una doble (un par de estrellas estrechamente enlazadas).

Es fácil calcular a mano la interacción gravitatoria entre dos estrellas o planetas. Las ecuaciones del movimiento se tornan, sin embargo, intratables para interacciones entre tres cuerpos. ¿Qué ocurre, por ejemplo, cuando una estrella solitaria colisiona con otra doble? Mientras la primera se mantiene suficientemente alejada de la doble, las dos que componen la estrella doble giran una alrededor de otra, en una órbita elíptica no perturbada. Si la estrella solitaria penetra en el sistema constituido por la estrella doble, el encuentro puede complicarse mucho más.

El resultado final de tal encuentro puede clasificarse en tres tipos cualitativamente distintos. Del choque pueden salir tres estrellas solitarias, no ligadas. En tal caso, como en física atómica, decimos que la estrella doble se ha ionizado. O bien originarse una estrella doble y otra solitaria. Si la estrella doble resultante está constituida por la intrusa y uno de los miembros de la estrella doble original, decimos que se ha producido una reacción de intercambio. Si los miembros de la estrella doble resultante son los componentes de la doble original, clasificamos el suceso como una aproximación. A veces, durante una colisión casi frontal, la estrella doble capturará temporalmente a la estrella solitaria. El sistema de tres cuerpos resultante puede permanecer enlazado durante largo tiempo —no son raras las relaciones inestables que duran centenares e incluso millares de períodos orbitales— pero, en última instancia, el sistema volverá a estar constituido por una estrella solitaria y otra doble.

Ese tipo de resultados es muy sensible a las condiciones iniciales. Basta un ligero cambio de la velocidad de la colisión o de los ángulos de orientación para alterar drásticamente la complicada danza de los tres cuerpos. No tiene sentido plantearse la posibilidad de un tratamiento analítico de tal comportamiento.

El número de parámetros necesarios para describir las condiciones iniciales de tales experimentos de dispersión es de nueve, cifra sobradamente excesiva para permitir un estudio sistemático de todas las combinaciones posibles. La dependencia del resultado experimental respecto de los valores precisos de los parámetros es extremadamente compleja. Aun manteniéndose fijos siete parámetros y variando sólo dos de ellos se obtiene una imagen nada sencilla [*véase la figura 3*].

Pueden aventurarse inferencias cuantitativas sobre los procesos de dispersión de tres cuerpos simulando un gran número de experimentos de ese tipo y efectuando posteriormente un análisis estadístico de los resultados. Los valores de la mayoría de los parámetros se escogen al azar, bajo restricciones que aseguren que los experimentos se desarrollen dentro del campo de interés de la aplicación astrofísica considerada. Cuando se realiza un número suficiente de experimentos, las incertidumbres introducidas por este método de muestreo de las condiciones iniciales, denominado de "Monte Carlo", son muy pequeñas y no impiden

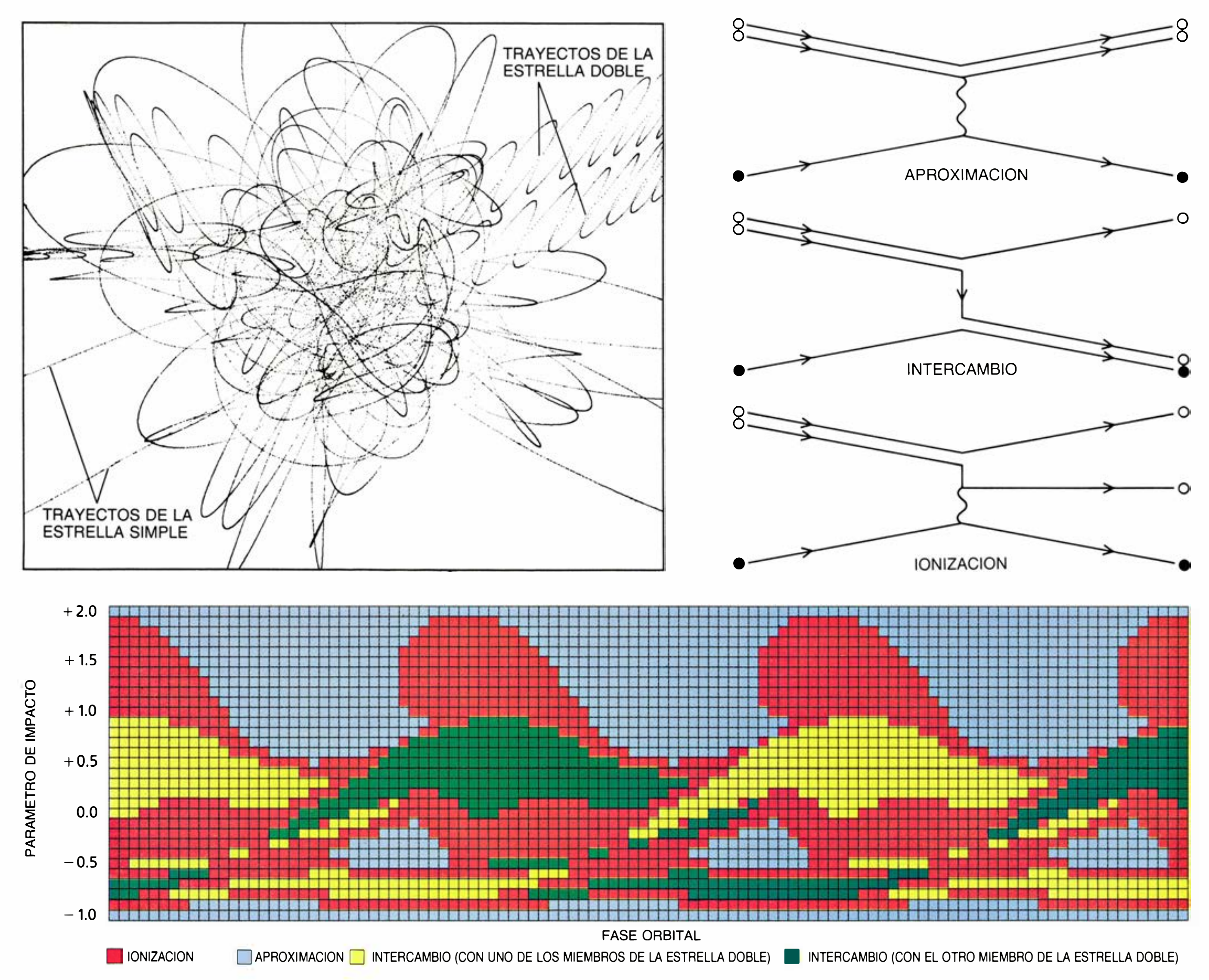
extraer conclusiones interesantes sobre el resultado promedio de los experimentos de dispersión de tres cuerpos.

¿Por qué interesa a los astrónomos esta clase de colisión? La respuesta reside en el papel que desempeñan las estrellas dobles en la generación de "calor". En una colisión entre una estrella doble y otra solitaria, la doble puede contraerse y transferir energía a la estrella solitaria, caldeándose el grupo de estrellas que las rodea. El proceso resulta análogo al de la fusión nuclear, en el que los núcleos chocan y se fusionan formando núcleos más pesados y liberando energía. La fusión nuclear es el mismo fenómeno que da brillo a las estrellas, incluido el sol.

Similarmente, la contracción orbital de estrellas dobles inducida por los choques puede caldear el núcleo de los cúmulos estelares densos, calor que compensaría la continua pérdida térmica superficial que registran dichos cúmulos. Para construir modelos de los cúmulos globulares debe conocerse el comportamiento estadístico de las colisiones entre estrellas solitarias y dobles. Del mismo modo que los ritmos de reacción nuclear constituyen un ingrediente esencial del cálculo de la estructura y la evolución de las distintas estrellas, las velocidades de reacción gravitatoria estelar son esenciales en el cálculo de la estructura y la evolución de los cúmulos.

El problema de la determinación de las velocidades de reacción gravitatoria es similar al que se le plantea al físico nuclear cuando se le pide que determine la velocidad de una reacción nuclear. Se puede dirigir un haz de núcleos de alta velocidad de determinado tipo hacia una placa que contiene núcleos de otro tipo. Contando el número de interacciones, midiendo sus propiedades y analizando estadísticamente los datos, el físico puede describir el comportamiento medio durante la colisión entre dos núcleos.

Un laboratorio informático permite un cálculo similar de las velocidades de reacción gravitatoria. Uno de los autores (Hut) ha participado en un pro-



3. CHOQUE entre una estrella solitaria y otra doble (un par de estrellas fuertemente enlazadas). Se trata de un acontecimiento complejo, como muestra la simulación por ordenador representada en la parte superior izquierda. Resultaría inútil tratar de calcular a mano ese fenómeno de dispersión de tres cuerpos. Por fortuna, se puede destilar la esencia del proceso haciendo caso omiso de los pormenores y clasificando únicamente el posible resultado final, igual que se estudia la dispersión de partículas en física subatómica. En este caso son tres los posibles resultados finales (*arriba, a la derecha*). En una aproximación, la estrella doble permanece intacta, si bien pueden resultar afectados sus parámetros orbitales. En un intercambio, el intruso remplaza a uno de los miembros originales de la estrella doble. En la ionización, las tres estrellas se independizan. Para describir las condiciones iniciales de estos experimentos de dispersión se requieren nueve parámetros distintos, excesivos para abordar un estudio sistemático de todas las combinaciones posibles. En el gráfico inferior, que resume el resultado de más de 1000 experimentos de dispersión, siete de los parámetros se mantuvieron constantes y se variaron sólo dos: el parámetro de impacto (el punto de mayor acercamiento de la estrella solitaria a la doble) y la fase orbital (las posiciones relativas que ocupan los miembros de la estrella doble).

yecto de investigación de la dinámica de los cúmulos estelares mediante la descripción estadística de la dispersión de tres cuerpos. Parte del proyecto consistió en realizar millones de experimentos de Monte Carlo de dispersión de tres cuerpos, en los cuales se mantuvieron fijos dos parámetros y se variaron siete de ellos, en un rastreo estadístico exhaustivo.

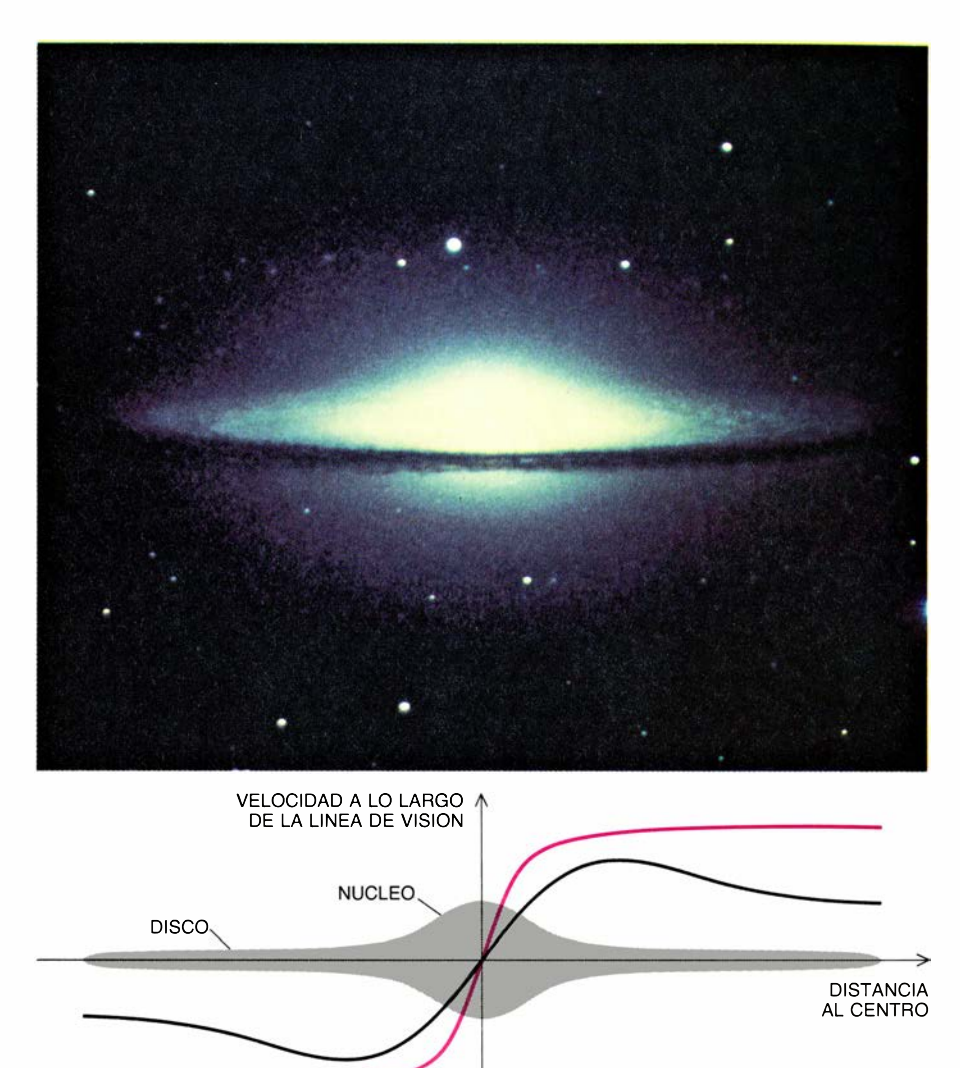
El experimento de dispersión de tres cuerpos quedaba al alcance de los recursos informáticos disponibles; requirió menos de un año de tiempo de ordenador en un VAX 11/780 de la Digital Equipment Corporation. Desarrollar un experimento similar con galaxias resulta mucho más difícil. Igual que los físicos de partículas provocan los choques violentos entre átomos y partículas subatómicas y examinan la "metralla" resultante para recabar información sobre la estructura interna de las partículas, los accidentes de tráfico galácticos proporcionarían datos muy importantes sobre qué clase de materia impregna el universo y cómo se distribuye. Avergüenza un poco reconocer que, si bien en las últimas décadas se ha avanzado bastante en el conocimiento del universo, no sabemos todavía de qué está constituido en buena parte. Las estrellas, el gas y el polvo —todo lo visible, con la simple mirada o con la ayuda de radiotelescopios y telescopios de infrarrojos, de rayos X y de rayos gamma— constituye sólo una pequeña fracción del total de la materia que aloja. La naturaleza de la materia invisible sigue siendo uno de los problemas pendientes de la astrofísica.

Se dispone de pruebas dinámicas valiosas de que la mayoría de galaxias encierran grandes cantidades de materia oculta. En las galaxias en forma de disco, las estrellas giran alrededor de un centro común. Según la ley de la gravitación de Newton, la velocidad orbital de una estrella en el disco debería descender rápidamente al aumentar la distancia al centro, si la única materia que mediara entre la estrella y el centro fuese la materia visible. En realidad, mediciones muy precisas indican que la velocidad rotacional no disminuye al crecer la distancia; antes bien, permanece casi constante a lo largo de una distancia considerable. La constancia de la velocidad de rotación implica que la mayor fracción de la masa no es visible y que ésa materia invisible se distribuye por un vasto volumen alrededor de las formas visibles.

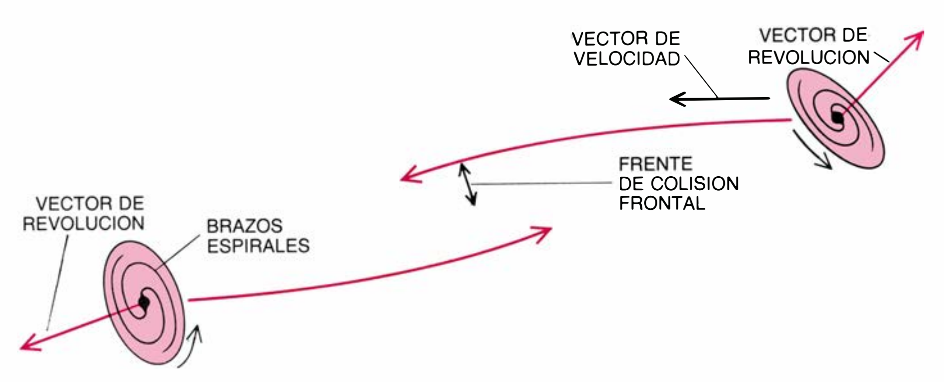
Las colisiones galácticas brindan la oportunidad de estudiar la distribución de la materia oculta. Los patrones producidos por las "salpicaduras" de la colisión revelan el campo gravitatorio de la materia invisible. En efecto, para seguir la pista de lo que de otra forma hubiese permanecido invisible, se emplean las estrellas como trazadores. El problema de la interpretación de sus trayectorias puede atacarse mediante una simulación masiva por ordenador. En este caso, se desconoce la distribución y la cantidad de masa de la galaxia. Con suficiente potencia de ordenador y soporte lógico, pueden simularse muchas colisiones con distintos parámetros y distribuciones de masa, buscándose los patrones de salpicadura de mayor parecido con los patrones realmente observados. Con unos modelos suficientemente sensibles, llegaremos a deducir la cantidad de masa y su distribución.

Una galaxia típica viene a constar de unos 100.000 millones de estrellas, cada una de las cuales interactúa con todas las demás mediante la fuerza gravitatoria. El cálculo de las interacciones entre tantas partículas queda fuera de las posibilidades de cualquier ordenador. En cambio, podemos obtener una aproximación al comportamiento de tal sistema mediante otro formado por muchas menos partículas: cada millón de partículas de la galaxia real equivaldría, en el modelo, a una sola.

¿Cuál es, en un experimento de colisión de dos galaxias espirales, el nú-



4. MATERIA OCULTA, o invisible que se encuentra en el universo, cuya existencia se infiere a partir de argumentos dinámicos. En una galaxia espiral típica, como la "del sombrero", o M104, las estrellas giran en torno a un núcleo central (*arriba*). El núcleo está rodeado por un disco (*abajo*) y un halo invisible. Según la ley de la gravedad de Newton, si entre una estrella del disco y el centro galáctico no hubiera más materia que la visible, la velocidad de rotación de las estrellas decaería al aumentar la distancia al núcleo (*curva negra de abajo*). La verdad es que la velocidad rotacional no disminuye, tal como indican las minuciosas mediciones realizadas; antes bien, permanece constante a lo largo de una gran distancia (*curva de color*). La constancia de la velocidad supone que en el gran espacio que queda entre las formas visibles existe materia invisible.



5. DISTRIBUCION DE LA MATERIA OCULTA en el interior de las galaxias. Se dedujo mediante simulaciones sistemáticas por ordenador de las colisiones intergalácticas. Con corrientes de estrellas se muestrea el campo gravitatorio de la materia invisible y se revela lo que de otro modo pasaría inadvertido.

mero mínimo de partículas necesario para que el modelo adquiera suficiente precisión y permita una buena comparación cuantitativa con el mundo real? Podemos representarnos una galaxia real como una densa masa central de estrellas, quizás algo aplanada por la rotación, un disco de estrellas y gas en órbitas predominantemente circulares y un halo de materia oscura más extenso, casi esférico, cuya existencia se infiere de razones dinámicas. Los componentes se mantienen en equilibrio estable, estacionario, en sus mutuos campos gravitatorios. En las galaxias espirales, el diámetro del disco centuplica, aproximadamente, su grosor. Deben, por tanto, emplearse al menos algunos centenares de miles de partículas para representar el disco; una cantidad menor no permitiría la resolución suficiente para representar el grosor finito del disco. Si se dota a todas las galaxias en colisión de un número comparable de partículas, cualquier experimento realista de dispersión deberá habérselas con casi un millón de partículas.

La integración de las ecuaciones de movimiento de un sistema constituido por un número de partículas en interacción comprendido entre medio millón y un millón constituye un reto formidable, que requiere lo más granado de los algoritmos numéricos y los ordenadores más potentes. Pero eso es sólo parte del problema: el margen de valores de los parámetros que describen las posibles condiciones iniciales de cada simulación del sistema es enorme. Para montar un experimento deben escogerse las distribuciones de cada componente y ensamblar los modelos de las galaxias. También tienen que elegirse los valores de todos los parámetros iniciales de la colisión.

El análisis de los resultados constituye el aspecto más arduo de la simulación numérica. Es necesario que parte del análisis se efectúe coincidiendo con la simulación de cada choque intergaláctico, para determinar si ha progresado suficientemente el ensayo y concluirlo. Debe acudirse también al análisis para dilucidar qué condiciones iniciales habrá que aplicar seguidamente al objeto de obtener los resultados más interesantes. Por último, el análisis extrae también la información referente a la estructura cualitativa y los parámetros del modelo necesarios para caracterizar los resultados del conjunto de experimentos.

¿En qué punto del experimento de dispersión podemos darlo por finalizado? Supóngase que contemplamos una película del proceso de dispersión. Las galaxias se aproximan una a otra, chocan y se "salpican" mutuamente estrellas y, o bien se fusionan en una única galaxia que pronto se estabiliza, o bien se separan y siguen su camino, quizá dañadas. Transcurrido cierto tiempo, puede determinarse si las regiones centrales de las galaxias se fusionarán o se separarán. En ese momento, sólo queda esperar la medición del alcance y el tipo de daños producidos.

Acotar el término de un experimento de colisión de galaxias y clasificar el resultado final constituyen tareas de reconocimiento de patrones cualitativos, de automatización difícil, pero no imposible. El sistema inicial lo forman dos moles, cada una de ellas reducida gravitatoriamente a un núcleo pequeño y denso. A medida que avanza el proceso de colisión, los núcleos se van separando en órbitas independientes o acaban por fusionarse. Si ocurre lo último, pueden presentarse varias etapas, a lo largo de las cuales las órbitas decaen y dilatan hasta que quedan estrechamente unidas. Durante el proceso, algunas de las partículas de cada mole se independizarán; otras formarán pellas, flámulas y colas transitorias, unidas o independientes.

El nivel de reconocimiento exigido

6. DIGITAL ORRERY, ordenador especializado en llevar a cabo simulaciones de órbitas planetarias a gran velocidad y con elevada precisión. El dispositivo consta de un volumen aproximado de 28 litros de equipo electrónico y consume la energía de una bombilla de 150 watt. En la resolución de los problemas para los que se diseñó, el ordenador realiza unos 10 millones de operaciones en coma flotante por segundo, es decir, opera 60 veces más deprisa que un VAX 11/780. La arquitectura en paralelo del Orrery, bien adaptada a los problemas de mecánica orbital, constituye la clave de sus altas prestaciones. Se ha utilizado ya el ingenio para determinar las órbitas que han seguido los cinco planetas exteriores del sistema solar en los últimos 100 millones de años y las que se espera que habrán de seguir durante los próximos 100 millones, período 40 veces más largo que el de cualquier integración de alta precisión anterior de los movimientos de los planetas.

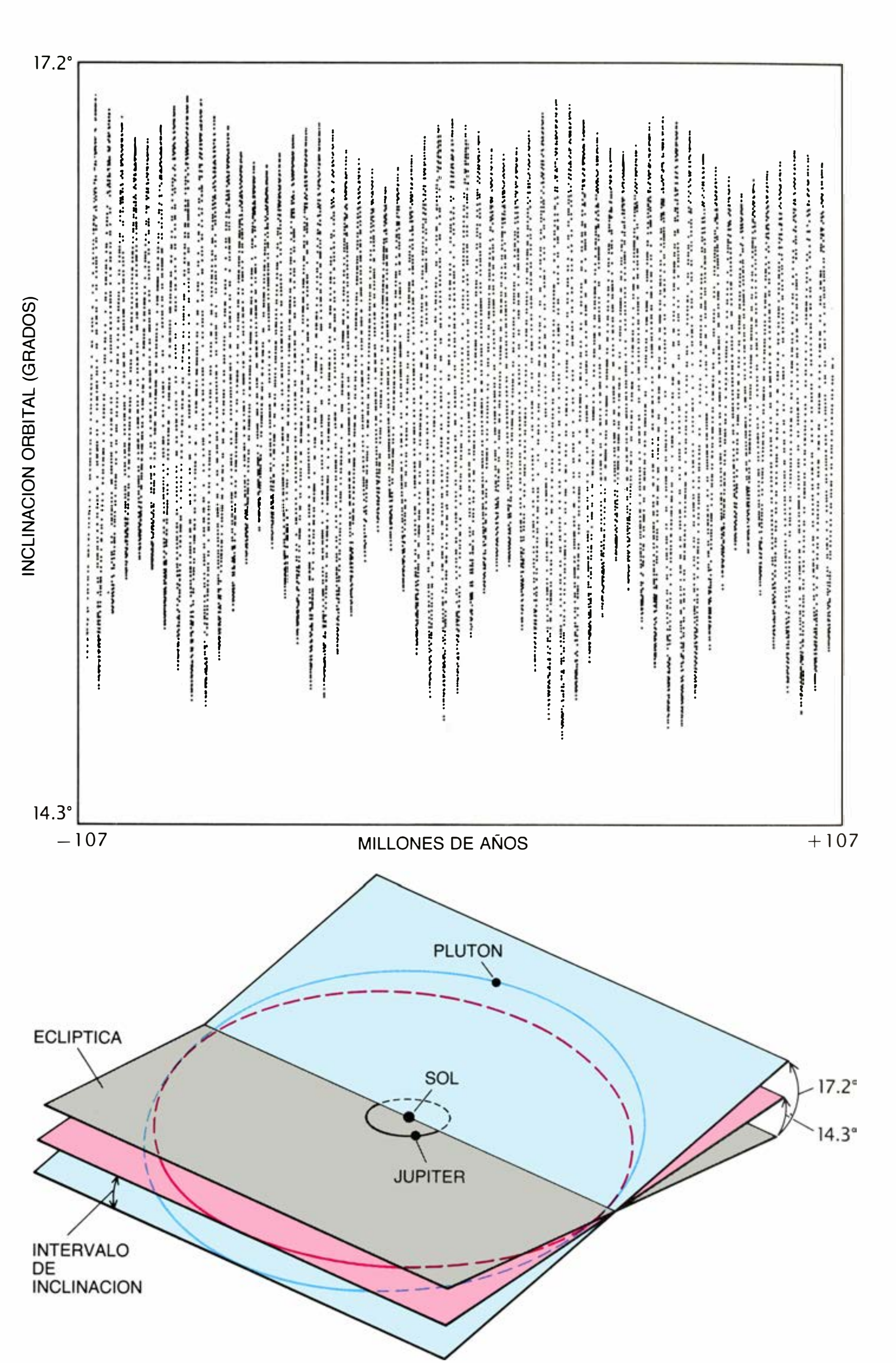
es bastante burdo. Hay que reseguir la evolución de un núcleo central que se mueve suavemente. No es difícil, en lo que al cómputo se refiere, hallar la posición de un núcleo central en un fotograma de película. Tampoco cuesta predecir la trayectoria a partir de las posiciones de los núcleos en instantáneas precedentes. Siguiendo de esa suerte la evolución de los núcleos, se determina si van a unirse o si se separarán.

Los autores se han embarcado en un proyecto a largo plazo de modelización de colisiones galácticas. Quizá tardemos varios años en obtener algún resultado definitivo, pues son ingentes los cálculos en juego. En nuestras simulaciones podríamos ayudarnos de procesadores paralelos [véase "Arquitecturas avanzadas de computadores", por Geoffrey C. Fox y Paul Messina, y "Programas para computación avanzada", por David Gelernter, ambos en este número].

Los cálculos científicos precisos para retrasar las fronteras del conocimiento requieren, a menudo, la aplicación de importantes recursos. Que ello sea así se debe, en parte, a que los modelos son de cómputo complejo y, en parte, a que la descripción de una clase de fenómenos obliga a realizar numerosos experimentos. Los científicos han accedido tradicionalmente a esos recursos adquiriendo grandes ordenadores o bien alquilándolos. Ambas opciones resultan caras. Los grandes ordenadores no sólo cuestan mucho dinero, sino que imponen también una pesada carga administrativa. Además, puesto que su estructura no suele adecuarse a cualquier problema, la redacción de programas idóneos puede convertirse en una tarea larga y compleja. Por supuesto, una parte sustancial del problema consiste en determinar exactamente qué algoritmos se requieren para investigar un fenómeno dado, pero la arquitectura del ordenador igual puede ayudar que estorbar en esa determinación. La redacción de los programas de hidrodinámica tridimensional que funcionen bien en un ordenador convencional demanda años de esfuerzo.

Un motivo importante del tamaño, costo y dificultad de programar un gran ordenador reside en la propia versatilidad de la máquina, diseñada para resolver problemas de muy diversa índole. Contrariamente, un ordenador especializado puede ser simple y de pequeño tamaño. A veces resulta incluso más fácil diseñar y construir la máquina que los programas. Es de esperar que los científicos y los ingenieros de ordenadores colaboren en el desarrollo de máquinas a la medida de sus demandas de cálculo. Podría así construirse un ordenador que resolviera determinadas ecuaciones diferenciales de física matemática en un tiempo similar, y con un gasto parecido, al requerido para preparar programas para esa aplicación en un superordenador. Además, el ordenador especializado puede convertirse en un instrumento experimental más, perteneciente al grupo que lo construyó, evitándose las trabas administrativas que caracterizan a los recursos compartidos o centralizados.

El diseño cooperativo de ordenadores especializados no se limitará a un puñado de máquinas que resuelvan



7. INCLINACION DE LA ORBITA DE PLUTON a lo largo de un período de 214 millones de años. Se ha determinado mediante el ordenador Digital Orrery (*arriba*). La inclinación de una órbita planetaria es el ángulo que forman el plano descrito por la órbita y la eclíptica, el plano que incluye la órbita de la Tierra (*abajo*). En la figura se aprecia una modulación de 34 millones de años de una oscilación fundamental de 3,8 millones de años; parece evidenciarse un ciclo de 137 millones de años. (Ilustración de Gabor Kiss.)

Científico: Consideremos la galaxia anular polar A0136-0801.

Ordenador: Dispongo de perfiles fotométricos y curvas de rotación para esta galaxia, tomados de Schweizer, Whitmore y Rubin (Astron. J., vol. 88, nº 7, julio de 1983). Aquí están. [Suponemos que los datos se representan gráficamente en pantalla].

Científico: Modelemos la distribución de masas de la galaxia como una combinación de un disco de Miyamoto con un halo isotérmico truncado. Primero ajustemos la medición del perfil fotométrico del huso al disco de Miyamoto.

Ordenador: No se produce un buen ajuste mínimo cuadrático. Esto es lo mejor que puedo hacer. [Nuevamente aparecería una imagen en pantalla.]

Científico: El problema está en que el anillo polar pasa sobre la parte noroeste del huso, donde decrece el brillo. ¿Observas que el perfil decrece en torno de −3 segundos de arco? Para realizar el ajuste deberías eliminar esos puntos anómalos de tu conjunto de datos.

Ordenador: De acuerdo. Puedo ajustar por mínimos cuadráticos el perfil fotométrico observado a un disco de Miyamoto a una escala longitudinal de 0,68 kiloparsec y una escala transversal de 0,92 kpc. Observa lo bueno que es el ajuste hasta que superamos los 10 segundos de arco.

Científico: Ahora, ajustemos la curva de rotación observada a la distribución de masas.

Ordenador: ¿Cómo debo usar los datos de curva de rotación para el anillo polar? Las velocidades circulares a radios iguales [el borde exterior del disco y el borde interior del anillo] parecen similares. ¿Crees que podemos fundir los conjuntos de datos?

Científico: Sí, intentémoslo de esa forma.

Ordenador: Ello me tomará cierto tiempo. Debo variar tres parámetros: la masa del disco del huso, la densidad central de la isoterma y el radio del núcleo de la isoterma. ¿Te importa esperar? Te llamaré cuando obtenga una respuesta adecuada.

[. . . Algún tiempo después.]

Ordenador: He encontrado una buena correspondencia. La masa del disco es de $1{,}75 \times 10^{10}$ masas solares. El halo isotérmico tiene una masa de $9{,}26 \times 10^{10}$ masas solares en 16,9 kiloparsec. El radio del núcleo en la isoterma es de 8,17 kiloparsec.

Científico: No entiendo por qué es tan grande el radio del núcleo . . . ¿cómo calculas la isoterma?

Ordenador: Mi potencial es proporcional a $\log(R^2 + B^2)$ para un radio del núcleo de B.

Científico: No me parece correcto; mi expresión del potencial contiene un arco tangente.

Ordenador: Entonces tu función de densidad es simplemente $\rho_0/(1 + R^2/B^2)$, ¿de acuerdo?

Científico: Sí.

Ordenador: Esto lo explica. Nuestras respectivas isotermas muestran el mismo comportamiento asintótico, pero medimos el radio del núcleo de distinto modo. En tus unidades el núcleo será unos tres kiloparsec menor.

Científico: Suena bien. ¿Qué razón masa/luz le calculamos al huso?

Ordenador: Cinco, aproximadamente.

8. DIALOGO HIPOTETICO entre un ordenador y un astrofísico, que ilustra la ayuda que podrían brindar en el futuro los sistemas interactivos a la hora de construir, simular y analizar modelos. Los autores prevén un sistema que puede mantener una conversación similar a la mostrada aquí, aunque probablemente en un lenguaje más formal. La característica más importante de este diálogo hipotético es el nivel conceptual de interacción con el ordenador: el programa muestra ser buen conocedor de la modelización de galaxias.

unos cuantos problemas. Tales proyectos renovarán profundamente el aire de la computación científica. Desde siempre, los científicos habían intervenido muy de cerca en el diseño de sus instrumentos, pero los ordenadores han recibido distinto trato. Los científicos son, ante todo, usuarios de la informática; sus necesidades las cubren unos cuantos fabricantes, muy alejados quizá, cuyo interés se centra prioritariamente en el mercado de las aplicaciones de gestión. Los científicos deberían participar en el diseño de sus propios instrumentos de computación. De hecho, el diseño de un ordenador constituye en gran medida un problema de ingeniería de programas, y la clave para que el diseño de ordenadores especializados forme parte habitual de la construcción de instrumental científico reside simplemente en convencerse de la viabilidad de ese empeño.

El cambio de actitud que prevemos se parece al que se registró hace unos años con motivo del diseño de circuitos integrados a muy gran escala (VLSI, por *very-large-scale integrated*), en los cuales se imprimen decenas de millares de componentes sobre una misma pastilla de semiconductor. En 1980, Carver A. Mead, del Instituto de Tecnología de California, y Lynn A. Conway, a la sazón en el centro de investigación que la Xerox tiene en Palo Alto, California, codificaron un conjunto de reglas de diseño para VLSI. Acabaron así con el misterio que rodeaba el diseño de VLSI y facilitaron el proyecto de microcircuitos especializados por parte de los propios diseñadores de ordenadores. El servicio de manufactura MOSIS, desarrollado a partir del trabajo de Mead y Conway, separa el diseño de un circuito de su fabricación. En la actualidad puede diseñarse un circuito y mandar las especificaciones a un servicio central de fabricación. Desde no hace mucho se aceptan las especificaciones del diseño de tarjetas de circuitos impresos, pudiéndose fabricar a distancia la mayoría de los componentes de un ordenador.

Ejemplo de ordenador científico especializado que se ha creado con dicha técnica es el Digital Orrery. Lo construyeron uno de los autores (Sussman) y un grupo de colaboradores. Se le diseñó para investigar la estabilidad a largo plazo del sistema solar. El ordenador está optimizado para la integración numérica de alta precisión de las ecuaciones de movimiento que describen un número pequeño de cuerpos, que ejercen una atracción gravitatoria mutua. Del diseño y construcción del Digital Orrery se encargaron tres físicos teóricos, dos informáticos y un técnico; lo terminaron en un año. El equipo electrónico del dispositivo ocupa un volumen inferior a los 30 litros y consume la potencia de una bombilla de 150 watt. En la resolución del problema específico para el que fue diseñado es 60 veces más rápido que un VAX 11/780 con acelerador de coma flotante. La máquina se emplea ahora en un estudio de las órbitas del cinturón de asteroides que pretende explicar la formación de los espacios vacíos denominados huecos de Kirkwood. El Digital Orrery ayuda también en la investigación de la estabilidad a largo plazo de la órbita de Plutón. Se ha usado para determinar las órbitas que supuestamente han seguido los planetas exteriores en los últimos 100 millones de años, y de las que seguirán durante los próximos 100 millones de años.

Antes de la construcción del Digital Orrery, la integración de alta precisión en simulaciones de períodos de millones de años resultaba prohibitivamente costosa. De ahí que sólo se realizaran unos pocos experimentos menores cuidadosamente programados. Pese a las obvias ventajas que brindan las máquinas especializadas, la comunidad astrofísica no había caído en la cuenta de diseñar y construir una. De hecho, se rechazó una primera propuesta de fabricación del Digital Orrery presentada a

la división de instrumentos astronómicos de la Fundación Nacional Científica de los Estados Unidos, alegándose que el proyecto resultaba inviable. Sorprendente tal actitud a la vista de los proyectos ingenieriles de construcción de telescopios, inmensamente complejos, que la misma comunidad ha llevado a buen término. Fabricar ordenadores no es tan difícil.

Mejorar la eficacia del cálculo científico exige superar los avances incrementales del soporte físico y de los programas. Debe mejorarse la eficacia del experimentador, proporcionándole el equivalente a una caja de herramientas, equipada con útiles que permitan construir y manipular modelos computacionales que capaciten para crear y controlar experimentos informáticos en función de los fenómenos cualitativos subyacentes. Las herramientas no deben limitarse a establecer los datos de una simulación numérica, ejecutarla y analizar los resultados; han de entender también cómo trabajan los modelos computacionales y cómo se relacionan con las demás descripciones del fenómeno estudiado.

Divisamos en el horizonte sistemas que combinarán potentes métodos numéricos con una manipulación simbólica compleja de descripciones tanto algebraicas como algorítmicas. Un entorno avanzado proporcionará apoyo simbólico-matemático para la preparación, ejecución y análisis de los resultados de experimentos numéricos. Los métodos simbólicos ayudarían a manipular las ecuaciones dinámicas, dándoles la forma adecuada para la simulación, y se emplearían para construir programas de simulación de alta calidad, optimizados para la situación particular estudiada, a partir de descripciones algorítmicas más generales. El científico debe ser capaz de discutir con el ordenador un modelo de galaxia del modo más parecido posible a como lo haría con un colaborador humano. Aunque no se haya construido aún ninguno, creemos que tales sistemas informáticos no son ninguna quimera.

A menudo, las herramientas nuevas han provocado en la civilización cambios no previstos por sus inventores; el procesamiento de datos ha superado ya el mero cálculo numérico. El cálculo simbólico y cualitativo está adquiriendo entidad propia en el entorno científico. Pasada la fase de meros masticadores de números, los ordenadores se convertirán muy pronto en insustituibles auxiliares de laboratorio.

Computación avanzada en el campo de la medicina

Con el tiempo, los ordenadores pueden llegar a ser tan imprescindibles como el estetoscopio. Los sistemas médicos pueden almacenar datos y recuperarlos selectivamente; no está lejos el día en que aconsejarán acerca del tratamiento

Glen D. Rennels y Edward H. Shortliffe

La señora Juárez padece un dolor en la parte baja de la espalda desde hace varias semanas. El reposo en cama y las aspirinas no han servido de ayuda. Supone que quizás un analgésico más potente le iría mejor y decide acudir al médico que le corresponde en la ciudad a la que se ha trasladado recientemente. Le explica al doctor que el dolor empezó al poco tiempo de haber jugado un partido de tenis por primera vez después de un año sin coger la raqueta. El examen físico sólo manifiesta cierta limitación de la movilidad en el codo izquierdo y un aumento de la sensibilidad en la zona baja de la espalda.

El doctor se dirige a su ordenador personal, que tiene acceso a una red nacional, y contacta con el ordenador hospitalario de la ciudad de donde procedía la señora Juárez. Ese ordenador contiene grabaciones médicas electrónicas a las que puede acceder el personal sanitario autorizado por el paciente. En pocos segundos, el médico comprueba que la señora Juárez fue sometida hace cinco años a una operación de tumorectomía para extirparle un pequeño cáncer de pecho. Desde entonces, no ha tenido ningún problema con su salud y había olvidado comentar el episodio al referirse a su dolor de espalda.

El doctor se desconecta de la red y contacta ahora con el QMR, un sistema de diagnóstico. El QMR, que posee conocimientos extensos acerca de las casi 600 enfermedades almacenadas en su memoria, sugiere que la señora Juárez puede tener artritis, tensión muscular o una recurrencia del cáncer con metástasis en la base de la columna vertebral. El médico, tranquilizado al comprobar que la "segunda opinión" del QMR coincide con su propio criterio, solicita una exploración ósea para descartar la posibilidad de una recurrencia del cáncer.

Ese desarrollo de los hechos no es ninguna fantasía del futuro. No está muy lejano el momento en el que los médicos considerarán al ordenador tan necesario como el estetoscopio. Los sistemas médicos de información que se están estudiando en los laboratorios de investigación pueden agruparse en dos categorías: sistemas de comunicación y sistemas de consejo. Los primeros almacenan la información médica, la recuperan selectivamente y la transmiten; los sistemas de consejo utilizan la información para ayudar a los médicos a diagnosticar la condición del paciente, como hizo el QMR en la situación descrita anteriormente, o para proponer, seguir y ayudar a controlar un tratamiento dado.

En algunos de los artículos precedentes que han abordado este mismo tema se describen los adelantos técnicos que aumentarán notablemente la eficacia de los sistemas. Aunque las mejoras del soporte físico ("hardware") redundarán en un mayor rendimiento de los sistemas de ordenador, ciertos laboratorios, incluido el nuestro de la Facultad de Medicina de la Universidad de Stanford, ponen su énfasis en el desarrollo de los métodos de soporte lógico ("software") capaces de representar la información y la experiencia médica en un aparato y de permitir que esa información y experiencia resulten accesibles para el médico. Así pues, podemos empezar por describir las capacidades presentes y futuras de los sistemas de comunicación y de consejo. Examinaremos ciertos problemas fundamentales en que siguen tropezando los técnicos de soporte lógico a la hora de crear programas que puedan "razonar" sobre un caso tanto como podría hacerlo un médico.

La necesidad de sistemas de comunicación surge en parte porque el médico (o el investigador bioquímico) encuentra cada vez mayor dificultad para leer, memorizar y recordar toda la información que requiere la resolución de un problema. Puede darse el caso de que el médico no se haya encontrado con la enfermedad o la queja específica de su paciente desde hace meses o años, intervalo durante el cual han podido cambiar las técnicas diagnósticas y terapéuticas.

Ante esa eventualidad, los médicos han empezado a acostumbrarse a recurrir a las bases de datos extraídos de la bibliografía pertinente. El MEDLINE es un ejemplo conocido. En la Biblioteca Nacional de Medicina de Bethesda, donde se desarrolló el sistema, un gran ordenador almacena casi todos los títulos de los artículos de la bibliografía mundial publicados durante los últimos 25 años y los resúmenes completos de los trabajos recientes. Cada artículo se indiza en palabras y frases clave, de forma que resulta fácil conseguir una lista de títulos relevantes y pedir después los resúmenes de artículos más apropiados. Existen sistemas de recuperación bibliográfica que almacenan el texto completo de los artículos y permiten localizar las frases pertinentes de cualquier documento.

Conforme se incorporen más prestaciones a su repertorio, el ordenador se convertirá en el principal medio de acceso a la bibliografía biomédica y a las bases de datos de enfermedades, siendo las bibliotecas las encargadas de todo ello. Estas, sin perder su acepción de espacio físico –edificios repletos de libros y revistas–, se considerarán fuentes suministradoras de información. No son, pues, dos conceptos excluyentes, sino que las nuevas tecnologías permiten que el modelo trascienda los límites que imponen las estructuras físicas.

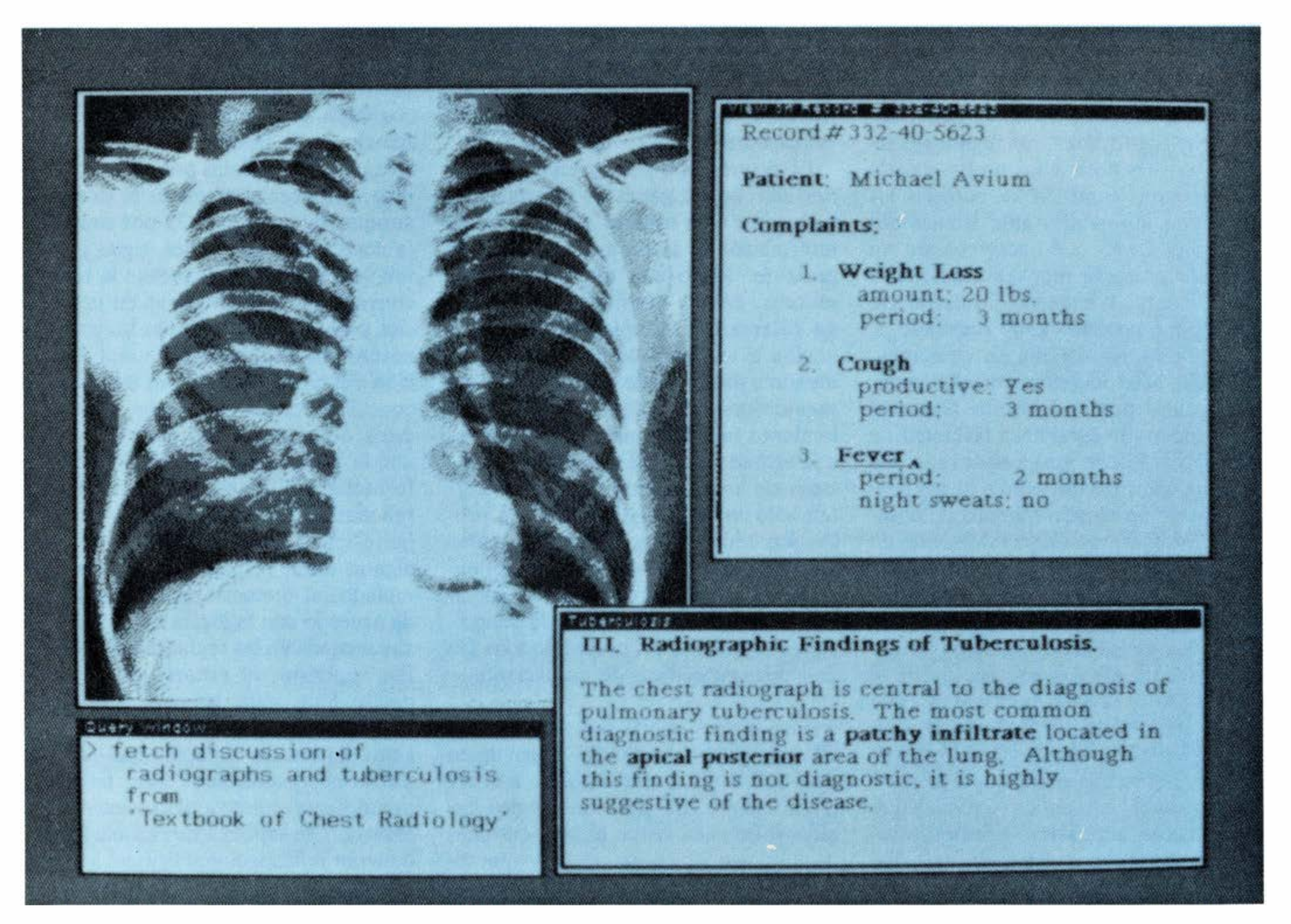
Los sistemas de comunicación posibilitan el acceso a la información; pero se encuentran también en desarrollo sistemas de consejo que puedan utilizar la información almacenada para solucionar problemas concretos del mundo real. En este orden, nosotros y nuestros colaboradores hemos venido trabajando en un sistema de consejo experimental: el ROUNDSMAN. Se trata de un sistema abierto en el que el médico expone un problema clínico y propone un plan de acción a seguir; acto seguido, ROUNDSMAN plantea una crítica a dicho plan.

Un objetivo central de este sistema es el de proporcionar acceso no sólo a los datos que se aportan en los artículos, sino también al *significado* de esos datos a la luz de la experiencia. En esta fase inicial, ROUNDSMAN puede elaborar análisis específicos, para cada paciente, de las diversas opciones de tratamiento del cáncer de pecho, basándose en 24 artículos. Para obtener esa capacidad, empezamos por codificar la información ofrecida por la bibliografía. Se almacenaron párrafos y puntos enteros; por ejemplo: "Doe y Roe NEJM 1986 ...en todos los casos incluidos en el estudio se extirparon los nódulos linfáticos y se confirmó la existencia del tumor... 473 casos... tratamiento aleatorio... 236 casos mediante mastectomía... 273 mediante tumorectomía ..." Se introdujeron también muchos otros detalles del plan de estudio y de los resultados.

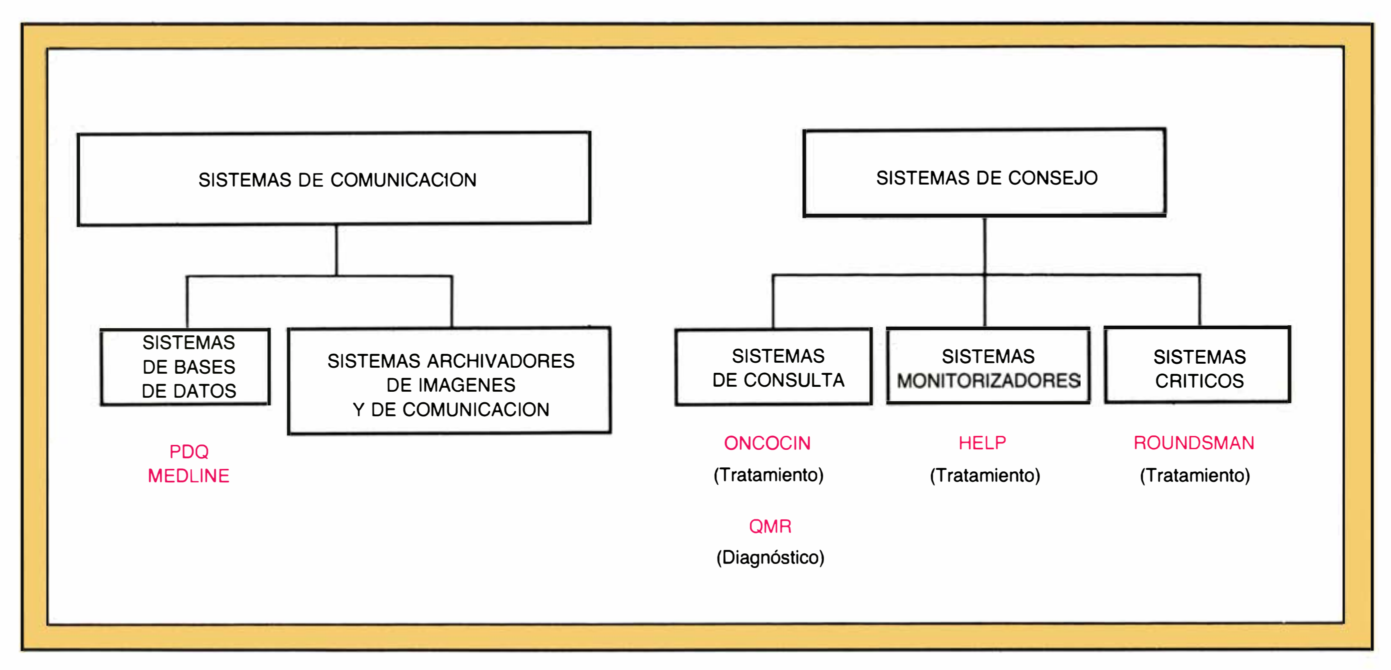
Solicitamos entonces de un experto (el tipo de médico al que consultaría un principiante sobre un caso difícil) su juicio subjetivo acerca de la información. Junto a los datos se almacenaron los suficientes juicios expertos como para cubrir las diferentes situaciones que podrían planteársele a ROUNDSMAN. En el supuesto de que el experto nos diga que las conclusiones del estudio deben aplicarse con cautela a los pacientes típicos de cáncer en estadio II, ya que los enfermos del citado artículo (que presentan nódulos linfáticos positivos) están más avanzados, el consejo del experto y su forma de abordar el problema se introducen en la memoria del sistema.

Imaginemos qué podría haber sucedido si la hipotética señora Juárez descubre por primera vez su tumor y acude a la consulta de un médico. Este la examina, recurre a ROUNDSMAN, le introduce las características de su paciente y propone practicarle una tumorectomía. ROUNDSMAN hubiera respondido: "Doe y Roe afirmaban en el *New England Journal of Medicine* en 1986 que la tumorectomía resultaba eficaz en tales casos, pero se debe señalar que la población del estudio se encontraba en un estadio pronóstico peor que el de su paciente". Tras hacer estas y otras consideraciones al médico, ROUNDSMAN hubiera sugerido que, pese a las discrepancias, los resultados de dicho trabajo aconsejan, en términos generales, la práctica de una tumorectomía.

La diferencia entre el soporte lógico para los sistemas de comunicación y el soporte lógico para los sistemas de consejo se pone de manifiesto al examinar las capacidades y los objetivos respectivos de dos sistemas experimentales, el Physician Data Query (PDQ) y



1. ESTACION DE TRABAJO MEDICA, mediante la cual la información de los departamentos de radiología, historias clínicas y biblioteca puede llegar hasta el despacho del médico interesado. En este ejemplo hipotético, el facultativo solicita la radiografía de tórax y la historia clínica de un tal Michael Avium. Los síntomas del paciente eran sugerentes de tuberculosis, de modo que el médico pidió al sistema que buscara en un manual una exposición suficiente acerca de las pruebas radiográficas de la tuberculosis (*abajo a la izquierda*). Acto seguido, el ordenador encontró y mostró la exposición al respecto (*abajo a la derecha*).



2. SISTEMAS INFORMATICOS MEDICOS, divisibles en sistemas de comunicación y sistemas de consejo. Los sistemas de comunicación almacenan, recuperan y transmiten el material bibliográfico, las historias clínicas de los enfermos (incluyendo los resultados de las últimas pruebas realizadas) y otros datos. Los sistemas de consejo participan activamente en el diagnóstico o en el tratamiento de los pacientes. Los sistemas en color se comentan en el texto.

el ONCOCIN. Ambos sistemas se refieren a protocolos de tratamiento del cáncer: son planes cuidadosamente especificados para llevar a cabo ensayos clínicos en los que se comparan las distintas opciones alternativas en un tratamiento. En el protocolo se exponen los procedimientos quirúrgicos o quimioterapeúticos particulares, los regímenes de radiación, las dosis y los programas de tratamiento. También se perfilan las conductas apropiadas ante situaciones delicadas, como qué hacer cuando un paciente presenta una reacción tóxica al tratamiento o cuándo considerar que este último no está dando resultado.

El PDQ es un sistema de comunicación: una base de datos controlada conjuntamente por el Instituto Nacional del Cáncer y la Biblioteca Nacional de Medicina. Puede proporcionar a los médicos información acerca de los protocolos de un amplio número de ensayos clínicos formales actualmente en curso de realización. En su ordenador personal, el médico puede acceder al ordenador central de la biblioteca y, merced a un soporte lógico especializado, identificar los protocolos recomendables para cada paciente. Característicamente, los protocolos presentan aproximaciones nuevas a los casos difíciles; aproximaciones que deben estar al alcance del médico (a menos que el protocolo se refiera a la administración de un fármaco experimental). De manera alternativa, el facultativo puede solicitar los nombres de los hospitales y de los médicos más cercanos que estén participando en un tratamiento determinado y remitirles a su paciente.

Supongamos que el examen óseo realizado a la señora Juárez confirma la presencia de un cáncer en la base de la columna vertebral. Su médico puede obtener del PDQ los nombres de los centros que trabajan en tratamientos avanzados de las metástasis del cáncer de pecho. Y, en consecuencia, aconsejar a su paciente la visita a un centro.

El sistema PDQ sólo pretende exponer una amplia gama de información, no ayudar a su usuario a aplicar dicha información al caso particular de cada paciente. Junto con nuestros colaboradores, hemos diseñado el ONCOCIN, un sistema que complementa al PDQ y auxilia a los médicos que están atendiendo a pacientes de cáncer tras haber seguido un protocolo y cuando el tratamiento ya está en marcha. ONCOCIN es a la vez un sistema de registro y un sistema de consejo (que todavía maneja tan sólo un pequeño número de protocolos terapéuticos). Vigila el curso del tratamiento y los progresos del enfermo, pero también es capaz de adquirir un papel activo e indicar la adaptación conveniente del protocolo a las necesidades específicas de un paciente.

Supongamos que la señora Juárez se incorpora a un ensayo clínico de un nuevo régimen farmacológico en el correspondiente hospital oncológico regional. En cada visita, el ONCOCIN puede servir de guía a los médicos que dirigen los complicados tratamientos quimioterapéuticos. Almacena los resultados de cada prueba, vigila el estado de la enferma y propone la manera de ajustar las dosis de cada droga en cada momento según la respuesta de la paciente al tratamiento.

Los detalles del protocolo se representan en el ordenador a través de unas estructuras conocidas por reglas de inferencia. Por ejemplo, una regla podría decir: "*si* el recuento de glóbulos blancos de la sangre del paciente está por debajo de lo normal, pero supera todavía las 3000 células por milímetro cúbico, *entonces* administrar la droga *A*, aunque dando sólo el 75 por ciento de la dosis habitual". Tales reglas permiten que el ONCOCIN aconseje la terapia conveniente a administrar en cada visita y que solicite también las pruebas necesarias. La pantalla de alta resolución del sistema simula un impreso, ya conocido por los oncólogos, en el que éstos describen a sus pacientes rellenando ciertas secciones en blanco del formulario. El ONCOCIN ofrece consejos rellenando los espacios vacíos de otras partes del impreso. Si el usuario tiene alguna duda respecto al consejo formulado, el ONCOCIN intenta explicarlo de acuerdo con la lógica del protocolo, almacenada en las reglas del programa. Por supuesto, el usuario es perfectamente libre para ignorar el consejo: ningún programa tiene acceso a la gama entera de variables que los médicos deben considerar antes de tomar una decisión relativa al cuidado de un enfermo. Se supone que el usuario puede tener información adicional a la que recurrir para ajustar el consejo del ONCOCIN.

Las características del soporte lógico (algunas de las cuales se comentarán más adelante) constituyen el principal

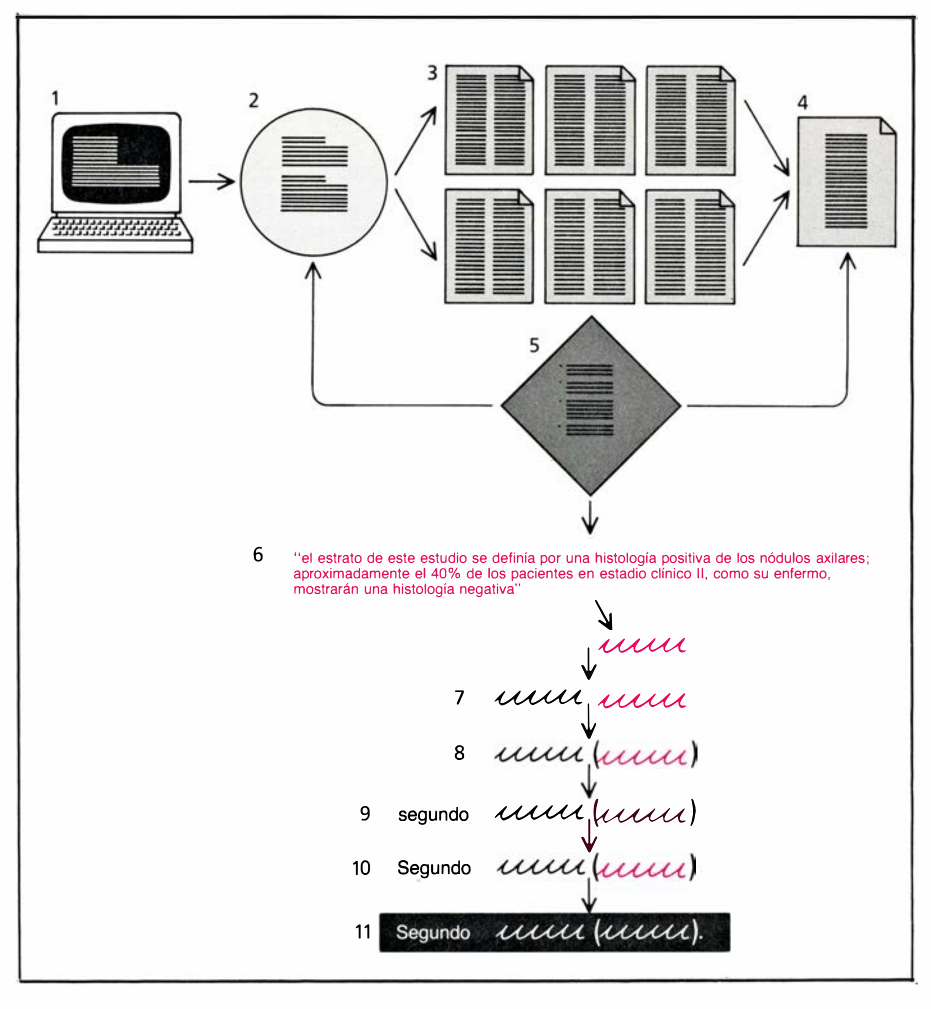
aspecto a desarrollar en el ROUNDSMAN, ONCOCIN y otros sistemas parecidos. Ahora bien, ciertos adelantos en el soporte físico, pensemos en procesadores más veloces o en pastillas de mayor memoria, facilitan, de hecho, el uso rutinario de esos sistemas. Hoy en día, por ejemplo, los circuitos de memoria pueden conservar sólo un número limitado de protocolos. Si el protocolo de un paciente no se encuentra en la memoria interna, la máquina debe localizar en un disco la información requerida e intercambiarla por la información que conserva en su memoria en ese momento. Se trata de un proceso lento, demasiado quizá para el uso rutinario en una clínica. Las pastillas de memoria de alta capacidad que se hallan en fase de desarrollo reducirán sustancialmente la cuantía de tal intercambio.

De estos comentarios acerca del ROUNDSMAN y del ONCOCIN se desprende la versatilidad de la interacción entre un sistema de consejo y sus usuarios. Como ya hemos mencionado, el ROUNDSMAN es un sistema crítico: reacciona frente a las conclusiones momentáneas del usuario —una propuesta de un plan de acción o una hipótesis diagnóstica— y sugiere, si es necesario, posibles alternativas. Por otro lado, un sistema de consulta proporciona análisis y recomendaciones acerca de un caso determinado independientemente de la opinión del usuario. Mientras que el ONCOCIN considera el tratamiento de la enfermedad del paciente, el sistema de consejo denominado QMR (de "Quick Medical Reference"), que ya mencionamos al comienzo de este artículo, hace hincapié en aspectos relacionados con el diagnóstico.

El QMR constituye una adaptación para microordenador del INTERNIST-1, un extenso programa diagnóstico desarrollado en la Facultad de Medicina de la Universidad de Pittsburgh. Maneja 577 enfermedades y sus interrelaciones con 4100 signos, síntomas y otras características del paciente. Los médicos pueden utilizarlo de tres formas distintas. En su modo básico, el QMR es un sistema experto de consulta que proporciona hipótesis diagnósticas específicas para cada paciente. En realidad, lo hace preguntándose a sí mismo con qué frecuencia los pacientes que muestran una característica determinada sufren cierta enfermedad y con qué frecuencia los individuos que padecen cierta enfermedad manifiestan una característica concreta. El sistema puede solicitar la información que necesita y que no se le había dado, así como sugerir pruebas y señalar puntos

```
investigators at multiple NSABP centers. Patients were
randomized to wide excision (& axillary dissection) and
adjuvant radiotherapy and adjuvant chemotherapy (N= 229) or
another protocol which was total mastectomy (& axillary
dissection) and adjuvant chemotherapy (N= 224). For patients
who underwent the first protocol the overall survival at five
years turned out to be 0.75 and recurrence-free survival at
five years was equal to 0.58. Under the second protocol the
overall survival at five years was 0.66 and recurrence-free
survival at five years was equal to 0.58.

How do these data apply to your patient? We are not
particularly concerned that the intervention was somewhat
nonstandard (they did not radiate supraclavicular nodes). More
troublesome is that first, there were modifications to one
intervention (in the excision arm, women with positive margins
received total mastectomy, but remained in the 'excision'
group). Second, the study population was in a worse prognostic
stratum compared to your patient (this study stratum was
defined by positive axillary node histology; about 40%
of clinical stage II patients like yours will have negative
histology).
```



3. ROUNDSMAN ofrece comentarios coherentes ante un plan propuesto de tratamiento, como en el caso del extracto, en el que se proporcionan datos sacados de un artículo publicado y el comentario al respecto (*arriba*). El proceso mediante el cual se ha generado la última frase se muestra en el recuadro inferior (*abajo*). El médico introduce una descripción de su paciente y propone un tratamiento (*1*). El sistema organiza esa entrada de información en un "contexto clínico" (*2*), busca un artículo pertinente (*3*) y extrae los datos de interés (*4*). Recurre a su base de conocimientos de juicios expertos (*5*) y la utiliza para contrastar el contexto clínico con los datos pertinentes. Destaca del texto una frase y la contrasta (*6*). El sistema antepone, a tal información, una generalización (*7*), introduce los detalles entre paréntesis (*8*), añade "segundo" para localizar la afirmación en una lista más extensa (*9*), pone las mayúsculas, puntúa (*10*) y ofrece el resultado (*11*).

Cover Sheet
Mass / X-ray
Disease Activity

Hematology	WBC x 1000	7.6	8.0	4.0	3.5	6.4	6.3	6.2	5.1	3.3	7.8
	% polys		54	24							
	% lymphs		20	30							
	PCV	32.8	33.9	27.2	27.3	27.6	25.7	26.5	24.3	30.6	
	Hemoglobin	11.1	11.4	9.1	9.4	9.5	8.8	9	8.2	10.4	
	Platelets x 1000	300	244	296	294	42	61	141	323	241	250
	Sed. Rate										
CHEMOTHERAPY (includes non-cytoxic drugs)	BSA (m2)										
	Arm assignment										
	Combination Name	POCC	VAM	POCC	POCC	VAM	VAM	VAM	POCC	POCC	
	Cycle #	1	2	2	2	3	3	3	3	3	
	Subcycle	B		A	B				A	B	
	Visit type	TREAT	TREAT	TREAT	TREAT	DELAY	DELAY	TREAT	TREAT	TREAT	
	Procarbazine (100)(MG/M2)x8	200		200	200				200	200	
	Vincristine (1.5)(MG/M2)x1	2.0		1.5	2.0				2.0	2.0	
	Cytoxan (600)(MG/M2)x1	1300		1300	1300				1300	1300	
	CCNU (60)(MG/M2)x1			130	0				130		
	VP16 (75)(MG/M2)x3		170					130			
	Adriamycin (50)(MG/M2)x1		110					80			
	Methotrexate (30)(MG/M2)x1		65					45			
	Cum. Adriamycin		230.0					310.0			

Radiotherapy
Symptom Review
Toxicity ?
Physical Examination
Chemistry ?
To order: Labs and Procedures
To order: Nuclear Medicine and Tomography
Sched:ling

Time	Day	6	27	20	31	24	1	8	29	5	26
	Month	Feb	Feb	Mar	Mar	Apr	May	May	May	Jun	Jun
	Year	87	87	87	87	87	87	87	87	87	87

Cover Sheet
Mass / X-ray
Disease Activity

Hematology	WBC x 1000	7.6	8.0	4.0	3.5	6.4	6.3	6.2	5.1	3.3	7.8
	% polys		54	24							
	% lymphs		20	30							
	PCV	32.8	33.9	27.2	27.3	27.6	25.7	26.5	24.3	30.6	
	Hemoglobin	11.1	11.4	9.1	9.4	9.5	8.8	9	8.2	10.4	
	Platelets x 1000	300	244	296	294	42	61	141	323	241	250
	Sed. Rate										
CHEMOTHERAPY (includes non-cytoxic drugs)	BSA (m2)										
	Arm assignment										
	Combination Name	POCC	VAM	POCC	POCC	VAM	VAM	VAM	POCC	POCC	VAM
	Cycle #	1	2	2	2	3	3	3	3	3	4
	Subcycle	B		A	B				A	B	
	Visit type	TREAT	TREAT	TREAT	TREAT	DELAY	DELAY	TREAT	TREAT	TREAT	TREAT
	Procarbazine 100MG/M2x14	200		200	200				200	200	
	Vincristine 1.5MG/M2x1	2.0		1.5	2.0				2.0	2.0	
	Cytoxan 600MG/M2x1	1300		1300	1300				1300	1300	
	CCNU 60MG/M2x1			130	0				130		
	VP16 (75)(MG/M2)x3		170					130			170
	Adriamycin (50)(MG/M2)x1		110					80			110
	Methotrexate (30)(MG/M2)x1		65					45			65
	Cum. Adriamycin		230.0					310.0			420

Radiotherapy
Symptom Review
Toxicity
Physical Examination
Chemistry
To order: Labs and Procedures
To order: Nuclear Medicine and Tomography
Scheduling

Time	Day	6	27	20	31	24	1	8	29	5	26
	Month	Feb	Feb	Mar	Mar	Apr	May	May	May	Jun	Jun
	Year	87	87	87	87	87	87	87	87	87	87

4. ONCOCIN consulta acerca de la quimioterapia por medio de una hoja de informe electrónica, mostrándose aquí la "imagen" de dos de ellas: una en un nivel de recogida de datos y otra en un nivel de consejo. Las ventanas para la hematología y la quimioterapia están abiertas; hay otras disponibles. El médico introduce los datos hematológicos del 26 de junio (*arriba*). Trabajando con esos datos y con la información del protocolo que se pretende aplicar, de la condición del paciente y del tratamiento que se ha dado hasta entonces, el ONCOCIN recurre a ciertas reglas para recomendar las dosis que deben administrarse al enfermo en su visita del 26 de junio (*abajo*); bajo condiciones diferentes, el sistema podría haber modificado las dosis o haber aconsejado una suspensión temporal del tratamiento. (Ilustraciones realizadas por Cliff Wulfman.)

que requieren una mayor investigación. Analiza la información obtenida y propone el diagnóstico o diagnósticos más probables.

En segundo lugar, el QMR puede funcionar como un libro electrónico, listando los síntomas del paciente que acompañan a una enfermedad o, inversamente, indicando cuál de sus 577 enfermedades se puede asociar con una característica determinada. En tercer lugar, puede combinar síntomas o enfermedades entre sí y determinar las implicaciones consecuentes. Por dar un ejemplo: se pueden especificar dos problemas médicos sin relación aparente y obtener sugerencias acerca de cómo las enfermedades coexistentes podrían, bajo circunstancias apropiadas, desencadenar ambos problemas.

Se ha presentado el INTERNIST-1, predecesor del QMR, como si fuera capaz de actuar casi con la misma eficacia que los especialistas en el diagnóstico de casos difíciles. El QMR tiene la misma base amplia de conocimientos y la misma capacidad diagnóstica, a las que se han añadido las funciones del libro de texto y de combinación de síntomas o enfermedades. Actualmente, el programa QMR está sometido a ensayo en Pittsburgh y en otras instituciones colaboradoras.

Cuando un sistema de consejo como el QMR se encuentra aislado, debe esperar las instrucciones del usuario para funcionar. Además, la persona que utiliza un sistema de este tipo debe introducir la información del paciente, que podría encontrarse ya en un sistema de comunicación y de manejo de datos cercano. ¿Podría un sistema de consejo acceder automáticamente a tal fuente de datos para prestar consejo siempre que fuera conveniente? Durante los últimos 15 años, se ha desarrollado en el Hospital LDS de Salt Lake uno de los primeros sistemas hospitalarios de información integrada capaces de cumplir ese propósito: el HELP.

El HELP es una extensión de un ordenador central que se conecta con las terminales e impresoras de todo el hospital. Por ejemplo, hay hoy, por lo menos, cuatro terminales y una impresora en cada servicio y una terminal en cada cama de la unidad de vigilancia intensiva; el proyecto es que exista una terminal junto a cada una de las 520 camas del hospital. El ordenador central es, ante todo, una instancia del manejo de datos, donde se almacena la historia de admisión de cada paciente y, siguiendo las medicaciones, los resultados de laboratorio y el estado actual de cada en-

fermo. Además de ser un sistema de manejo de datos, el ordenador central incorpora miles de subprogramas especiales lógicos: los sectores HELP. Tiene por misión revisar y vigilar todos los datos disponibles de cada paciente y mantenerse atento a una serie de condiciones de alerta especificadas. En cuanto identifica alguna de éstas, HELP manda un mensaje de aviso a los miembros del equipo pertinente.

Cuando se hospitalizó a la señora Juárez para operar su cáncer original, un sistema HELP podría haber controlado el proceso de medicación, sus resultados en las pruebas de laboratorio y sus constantes vitales. A diferencia del médico, este sistema puede revisar los resultados de las pruebas en segundos y durante las 24 horas del día. Supongamos que la señora Juárez estuviera tomando el fármaco digoxina por padecer una alteración cardiaca. Si los resultados de sus análisis de sangre revelaran un nivel de potasio bajo, el HELP habría avisado al enfermero de la planta advirtiéndole que la disminución de los niveles de potasio constituye un peligro para un enfermo que está tomando digoxina y que la señora Juárez podría necesitar de inmediato la administración de una solución oral de cloruro potásico.

HELP es un ejemplo ideal de cómo aumentar la flexibilidad y el rendimiento de un sistema de consejo mediante su integración en un sistema de comunicación. Y HELP es sólo el principio. Los avances en las técnicas de redes de área local permitirán incorporar a los sistemas de información hospitalarios un sistema de consejo especializado como el ROUNDSMAN, un sistema diagnóstico e incluso un sistema que facilite la llegada de las radiografías, las ecografías y otros medios diagnósticos por imagen desde el departamento de radiología hasta una terminal en cada quirófano y junto a cada cama.

Lista de asociaciones

Enfermedad pulmonar y DIARREA crónica

Parejas de enfermedades coherentes con los hallazgos y los temas introducidos

Atelectasia
causada por un síndrome carcinoide secundario a una neoplasia bronquial

Derrame pleural exudativo
causado por un pseudoquiste pancreático

Fibrosis intersticial pulmonar difusa secundaria
causada por una esclerosis sistémica progresiva

Hipertensión pulmonar secundaria
causada por una esclerosis sistemática progresiva
o simultánea a una esquistosomiasis hepática crónica

Infarto pulmonar
predispuesto por un carcinoma del cuerpo o de la cola del páncreas
o predispuesto por un carcinoma de la cabeza del páncreas
o causado por una obstrucción de la vena hepática

Legionelosis pulmonar
predispuesta por un síndrome de inmunodeficiencia adquirida <SIDA>

Linfoma pulmonar
coincidente con un linfoma de colon
o coincidente con un linfoma de intestino delgado

Neumonía eosinofílica aguda <LOEFFLER>
causada por la enfermedad de Hookworm

Neumonía neumocócica
predispuesta por un síndrome carcinoide secundario a una neoplasia bronquial

Neumonía por Pneumocystis
predispuesta por un síndrome de inmunodeficiencia adquirida <SIDA>

5. SE LE PIDE A QMR (*color*) una lista de enfermedades que podrían justificar la presencia de dos problemas aparentemente dispares: la enfermedad pulmonar y la diarrea crónica. La lista que aquí se expone es una parte de la respuesta completa del QMR, generada por el sistema en función de su base de conocimientos, en la que almacena 577 enfermedades, 4100 características de los enfermos y sus interrelaciones.

Si tal es su poder y acceso en versiones experimentales, ¿por qué los sistemas médicos avanzados no gozan de un uso generalizado en la actividad clínica diaria? Lo cierto es que los métodos de fabricación de los sistemas de consejo no están maduros: en el diseño de un soporte lógico efectivo todavía se plantean algunos problemas teóricos y prácticos.

Aspecto fundamental en la mayoría de sistemas de consejo es el relativo al diseño de determinadas habilidades clínicas, especialmente del proceso de tomar de decisiones. ¿Cómo se diseña el modelo de tal capacidad especializada y eminentemente humana? Existen muchas estrategias basadas en aproximaciones tan dispares como el reconocimiento de patrones, la regresión estadística, los algoritmos dendriformes y la teoría de la decisión. El análisis de otro enfoque, el de la inteligencia artificial (IA), nos mostrará algunas de las dificultades que se encuentran en el desarrollo del soporte lógico de los sistemas de consejo.

Los investigadores de la IA médica han recibido influencias de las corrientes psicológicas que sugieren que el principal factor que caracteriza a los médicos expertos es la riqueza de su conocimiento —es decir, su acopio de información acerca de enfermedades, síntomas y tratamientos—, más que la forma particular de procesar esa información. Los programas de ordenador utilizados en las técnicas de IA para representar y manipular los conocimientos de un determinado campo de especialización se denominan sistemas expertos. Para desarrollar un sistema experto, los investigadores de la IA médica se dedican, en primer lugar, a buscar una representación óptima de la información médica en el ordenador. Gran parte de las ventajas de un sistema médico basado en los conocimientos deriva de su almacenamiento de información y de la forma en que ésta se estructura, más que de las nuevas técnicas de procesamiento (tales como un algoritmo que el ordenador manejara mejor que un ser humano).

Uno de los principales retos en la construcción de un sistema basado en la información es el de tratar la cantidad total de datos que deben representarse y manipularse. Tomemos, por ejemplo, las 577 enfermedades y las 4100 características de los pacientes que almacena el QMR. Cuando se relacionan y se valoran esas 4677 entidades (para mostrar que ciertos síntomas sugieren ciertas enfermedades, que una enfermedad puede ser causa de otra y que determinadas características de los pacientes son más importantes que otras), la base de información del QMR puede contener unos 250.000 conceptos. Muchos de ellos no están explícitamente recogidos en ningún libro de texto. Son necesarios 10 años de colaboración entre los especialistas en ordenadores y los médicos para conseguir que cada tema esté bien acotado y sea suficientemente explícito para su representación en el sistema. Además, el QMR abarca sólo una parte de las enfermedades más importantes conocidas. Por tanto, labor primordial en el diseño

REGLA
Para determinar la dosis de metotrexate administrado en la quimioterapia VAM en los protocolos 20-83-1 y 2091:
Si: el nivel de creatinina en suero (en mg por dl) es superior a 1,5
Entonces: no dar metotrexate

ESQUEMA		
Nombre: Tularemia		
Síntomas y signos:	Intensidad provocada	Frecuencia
FIEBRE	0	5
CULTIVO de Francisella tularensis de la lesión cutánea	5	4
NODULOS linfáticos aumentados	1	4
ULCERA cutánea	1	4
TAQUICARDIA sinusal en el ECG	0	4
PRUEBA cutánea de tularensis positiva	4	3
EXPOSICION a conejos, roedores o micromamíferos	2	3
PICADURA reciente de garrapata	2	3
CEFALALGIA severa	1	3
Puede producir:		
Shock pirogénico	1	2
Apendicitis aguda	1	1
Predispone a:		
Endocarditis infecciosa aguda en corazón izquierdo	1	1

6. REGLAS Y ESTRUCTURAS son los elementos de dos enfoques para representar el conocimiento médico. Las reglas tienen la forma de "*si* se da la condición *X*, *entonces* se actúa de la manera *Y*". Cuando el ONCOCIN, un sistema basado en las reglas (*arriba*), al determinar las dosis adecuadas para la quimioterapia VAM establece que la parte "si" de su regla del metotrexato se cumple, concluye que la parte "entonces" es la indicada. Las estructuras son detalles descriptivos de las enfermedades, los síntomas y otras entidades similares. Dentro del QMR, un sistema basado en las estructuras, en el cuadro nosológico para la tularemia se relacionan de forma ponderada varios signos, síntomas y otras enfermedades. El valor de la "intensidad evocada" refleja la probabilidad de que la enfermedad esté presente cuando se observa un determinado síntoma: el valor O significa que el síntoma es inespecífico y el valor 5 que la enfermedad es la única causa del síntoma. Por otra parte, en la "frecuencia" se señala la probabilidad de que la enfermedad origine un determinado síntoma: el 1 significa que es muy poco probable; el 5 indica que aparecerá casi siempre.

de sistemas expertos será la de identificar los elementos claves del conocimiento y sus interrelaciones y codificarlos en estructuras y conexiones de datos que permitan aplicar la información para solucionar los problemas.

Un segundo problema de diseño es el de asegurarse de que las estructuras de datos que representen los elementos de conocimiento permitirán que se amplíe y se corrija la información del sistema. En medicina, los descubrimientos modifican el acto clínico y, por tanto, un sistema basado en la información debe aceptar las correcciones y las ampliaciones de su base de conocimientos. Los remiendos sobre el sistema completo para introducir cada avance son inviables. En consecuencia, lo ideal es que la estructura de datos (sea una lista, una jerarquía, una red o cualquier otro tipo) que representa el conocimiento médico sea independiente de los algoritmos que manejan la información. Esa independencia resulta difícil de conseguir, ya que ciertos conocimientos médicos son, en sí mismos, algorítmicos en parte, por cuanto expresan una estrategia para resolver un problema más que un hecho real. Frecuentemente, por ejemplo, los diagnósticos pretenden establecer el arco general de una enfermedad —como la "enfermedad hepática"— antes que concretar una forma particular de hepatitis.

Por otro lado, cada elemento de conocimiento médico debería representarse de una forma modular, de manera que la alteración de un elemento no afectara a los demás. Dos relaciones de datos que pretenden satisfacer tal modulación son las reglas y las estructuras. Llamamos regla a un estatuto inferencial de la forma "si A, entonces B" por medio del cual un sistema puede alcanzar una conclusión a través de la norma lógica del *modus ponens*: si A implica B, y si A es cierto, entonces se puede inferir B. Otra regla puede utilizar los resultados inferidos por una regla anterior para hacer una segunda deducción: las reglas "se encadenan".

A su vez, una estructura representa una entidad, tal como una enfermedad, que se caracteriza por "muescas" que definen las peculiaridades prototípicas de esa entidad, verbigracia, los síntomas que acompañan a la enfermedad. También es posible definir los métodos para deducir los valores de las muescas a partir de otros aspectos de la estructura o de otras estructuras.

La modularidad que permiten las reglas y las estructuras es, de suyo, difícil de conseguir, entre otras razones porque el conocimiento médico es más que la suma de las partes. Por ejemplo, el ROUNDSMAN representa cada artículo de revista como un objeto diferenciado en su base de conocimientos. Cuando se publica un nuevo artículo, sería deseable aumentar la base de conocimientos de una forma modular. Sin embargo, la información que contiene un artículo arroja con frecuencia una nueva luz sobre los artículos anteriores, modificando su interpretación. Cuanto más se intenta acomodar esa peculiaridad no modular del conocimiento médico —añadiendo elementos de enlace entre los artículos para representar sus interrelaciones—, más se tiende a crear una base de conocimientos enmarañada, de elementos estrechamente dependientes, que resulta impenetrable cuando se trata de modernizarla.

Un tercer obstáculo en la investigación surge de la limitación de modelos existentes del cuerpo humano o de la enfermedad. Ello contrasta con los sistemas supuestamente diseñados para facilitar un proceso de manufactura. En el manufacturado hallamos una comprensión clara de la función de los distintos componentes y del engarce mutuo; los problemas de producción y los diagnósticos fallidos se pueden abordar tratando las características combinatorias de una secuencia de uniones o de las formas en que los componentes deben estar funcionando mal.

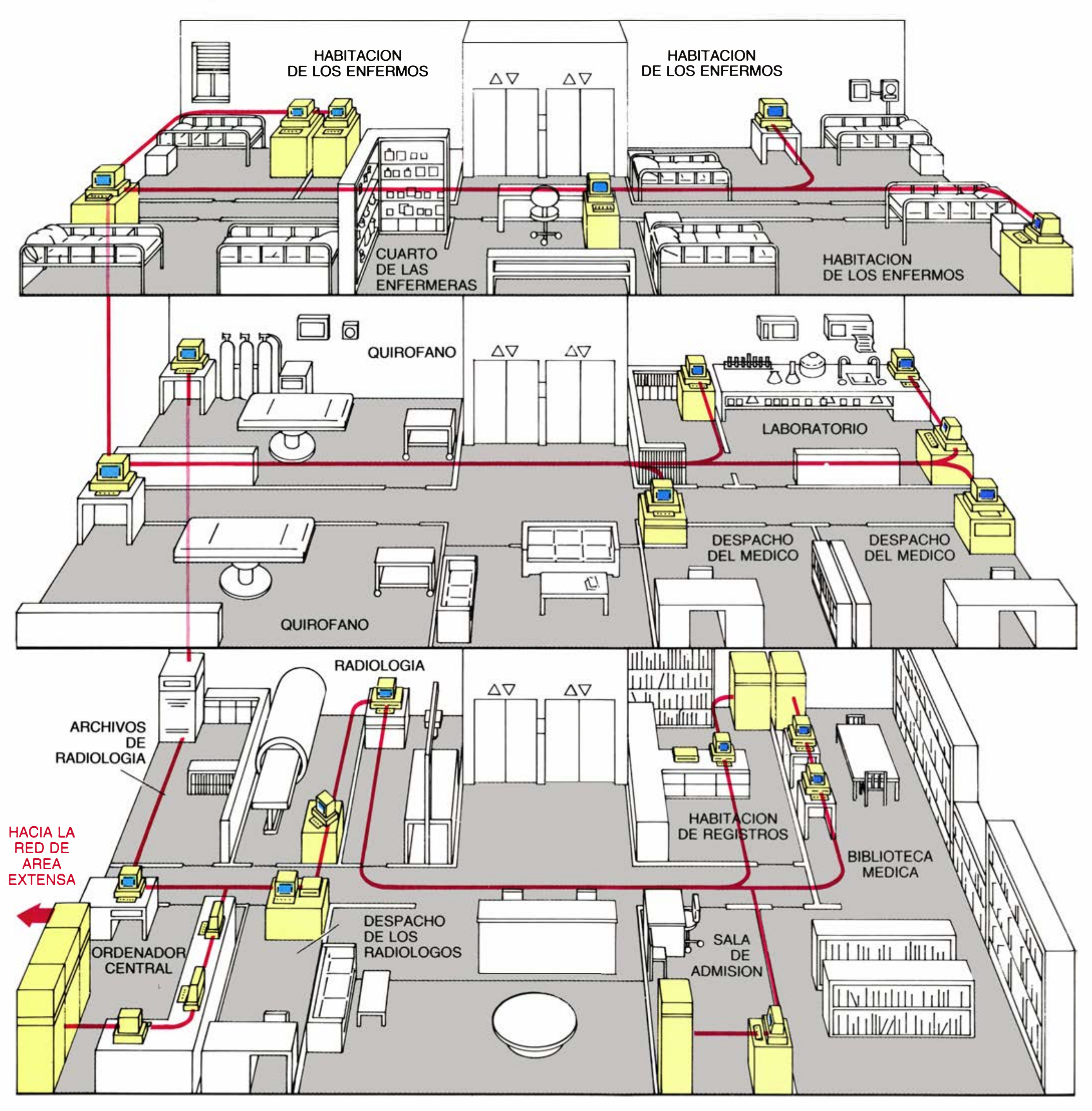
Pocos elementos y pocas funciones del cuerpo humano pueden comprenderse enteramente desde esa perspectiva mecanicista. Por tanto, los médicos necesitan hacer uso de una amplia gama de conocimientos, incluyendo reglas empíricas aprendidas en las mejores clínicas, experiencia derivada del análisis estadístico de ensayos clínicos, mecanismos causales deducidos de la anatomía, la fisiología y la biología molecular y aspectos sociales relacionados con la calidad de la asistencia médica. A los expertos en ordenadores médicos compete, pues, desarrollar las técnicas necesarias para codificar y representar esas formas diferentes de conocimiento de una manera coordinada.

El trabajo conceptual y de diseño que llevan a cabo las personas que desarrollan los sistemas de consejo se puede contrastar con los problemas prácticos que abordan quienes se ocupan de los sistemas de comunicación. Nos haremos una idea de esos problemas si consideramos los obstáculos que se encuentran al intentar llevar a la práctica un sistema archivador de imágenes y de comunicación denominado PACS. Tal sistema se halla en una fase mucho más avanzada que la propia de

investigación, pero aun no está disponible porque ciertas barreras logísticas y económicas impiden su aplicación práctica.

El PACS almacena imágenes de rayos X y de otros tipos en forma digital, las entrega cuando se necesitan y ofrece en la pantalla imágenes de alta calidad fotográfica. Los médicos tienen todavía que trasladarse hasta el departamento de radiología para ver las imágenes radiológicas; frecuentemente, los pacientes deben llevar sus placas de rayos X de un despacho a otro. Se ganaría muchísimo si las imágenes radiográficas pudieran distribuirse directamente a los hospitales y clínicas, en particular si la misma imagen pudiera contemplarse, a la vez, por médicos distintos en diferentes lugares.

El soporte lógico para producir imágenes y ofrecerlas de esa manera está ya preparado. Existen líneas de comunicación por fibra óptica que viabilizan la velocidad de señal necesaria, de 200 megabits por segundo, y se dispone de terminales que ofrecen las imágenes con la suficiente resolución. (De hecho, se están utilizando en las industrias del cine y de la televisión.) Si bien es verdad que los sistemas digitales de imagen se beneficiarían de una mayor investigación, la tecnología actual podría mejorar radicalmente la distribución y la interpretación de las imágenes médicas. Sin embargo, en el presente clima económico competitivo de la medicina resulta difícil justificar una tecnología nueva y cara.



7. SISTEMA DE INFORMATIZACION HOSPITALARIA, en una visión de futuro, que utilizará una red de difusión local para transmitir textos y también imágenes hasta las estaciones de trabajo de cualquier zona del hospital. La información de cada paciente se introducirá desde —y se transmitirá hacia— la sala de admisiones, la sala de historias clínicas, los laboratorios, los quirófanos y la cama donde se encuentre. Los resultados, almacenados digitalmente, de los rayos X, los escáneres, los ultrasonidos y otros exámenes se transmitirán por cables de fibra óptica hacia cualquier lugar en donde se necesiten. El material de la biblioteca también será accesible y una amplia red permitirá acercar las bases de datos y los sistemas de consejo. A diferencia de los sistemas convencionales, dirigidos desde un ordenador central, los sistemas del futuro descentralizarán en estaciones de trabajo individuales gran parte de la memoria y del procesado. Un sistema como el descrito sería una versión más compleja del sistema HELP, que ya se ha llevado a la práctica en el Hospital LDS de Salt Lake.

¿Qué supondría la introducción de un PACS en un hospital moderno y concurrido? La instalación de la red de fibra óptica sería complicada y cara. Para una distribución adecuada de las imágenes de alta resolución, el hospital necesitaría múltiples estaciones de trabajo, dotada cada una de ellas con una oferta de resolución de 2048 por 2048 píxeles (elementos discretos de imagen), siendo cada píxel capaz de registrar de 5 a 8 tonalidades de gris.

Ante el gasto que exigiría la instalación de tal sistema, algunas instituciones han ido adquiriendo poco a poco equipos de imagen digital, comprando una máquina digital cuando había que sustituir un aparato viejo. Así, poco a poco, van haciéndose con los suficientes equipos digitales y estaciones de trabajo para integrarlos en una red PACS. Pero esto es sólo un ejemplo de cómo los obstáculos económicos y logísticos pueden impedir la generalización de una tecnología aprovechable.

Hoy en día, los sistemas de ordenadores no tienen —ni tendrán a corto plazo— una comprensión suficiente de las técnicas, la clínica y los aspectos sociales de la medicina que se aproxime a la riqueza y la flexibilidad de la experiencia humana. Los sistemas de consejo que hemos descrito representan adelantos en las ciencias informáticas de la medicina, pero ninguno de ellos sirve sin el sentido común y el juicio del médico experimentado. Quedan muchos aspectos en la resolución de los problemas humanos que ni nosotros, ni otras personas de nuestro campo de trabajo, sabemos cómo modelar en el interior de un ordenador.

A pesar de esas cortapisas, nos sentimos optimistas acerca del futuro de la aplicación de los ordenadores avanzados en el área de la medicina. Los ordenadores no son meros devoradores de números. Su capacidad para almacenar, recuperar selectivamente y transmitir la información médica se halla fuera de toda duda. En el horizonte aparecen sistemas que proporcionarán consejos razonados sobre el diagnóstico y la resolución de casos específicos. Los adelantos en las ciencias materiales, en el diseño de circuitos, en el trazado de redes y en el procesado en paralelo prometen, todos ellos, una mayor libertad para quienes, como nosotros, trabajan en el soporte lógico de los sistemas médicos de información. Esperamos ver la realización y el uso creciente de sistemas avanzados que sean lo suficientemente "inteligentes" como para asumir un papel activo, aunque subordinado, en la praxis médica.

Computación avanzada en la industria

Los superordenadores podrían desempeñar un papel de primer orden en la industria. Han participado de modo destacado en el proyecto de diseños aerodinámicos de aviones y automóviles

Albert M. Erisman y Kenneth W. Neves

La disponibilidad de superordenadores capaces de realizar cuantiosas operaciones aritméticas con gran rápidez ha brindado a los diseñadores industriales nuevas perspectivas del análisis computacional. Con ello, las máquinas han abierto la vía hacia mejores productos y sustanciales ahorros en las fases de diseño y desarrollo de los mismos.

Sin embargo, y a pesar de esos logros, los superordenadores plantean aún numerosos problemas. En particular, su plena adopción por la industria aeronaútica y otros sectores constituye todavía un objetivo de difícil alcance. A modo de ejemplo, ciertos problemas de interés industrial resultan demasiado complejos para que incluso los computadores más avanzados puedan resolverlos. Por otra parte, las limitaciones de la programación actual restringen la aplicación de los superordenadores en la industria. Finalmente, queda por resolver la transición desde el análisis computacional en el superordenador hasta la plena integración del diseño asistido por ordenador (CAD) y la fabricación asistida por ordenador (CAM). La obtención de tal integración es una meta aún no alcanzada. Un repaso al diseño aeronáutico aclarará esos puntos.

El campo de la aerodinámica fue de los primeros en recurrir a los ordenadores para resolver problemas de cálculo. Hace ya 20 años, los ingenieros aeronaúticos los utilizaban para recrear el flujo del aire alrededor de un ala y, con muchas más limitaciones, el flujo alrededor del cuerpo del avión. Hoy, los superordenadores permiten a los especialistas en esos campos modelar aviones enteros y distinguir las diversas componentes del arrastre aerodinámico que se generan en cada sección del avión.

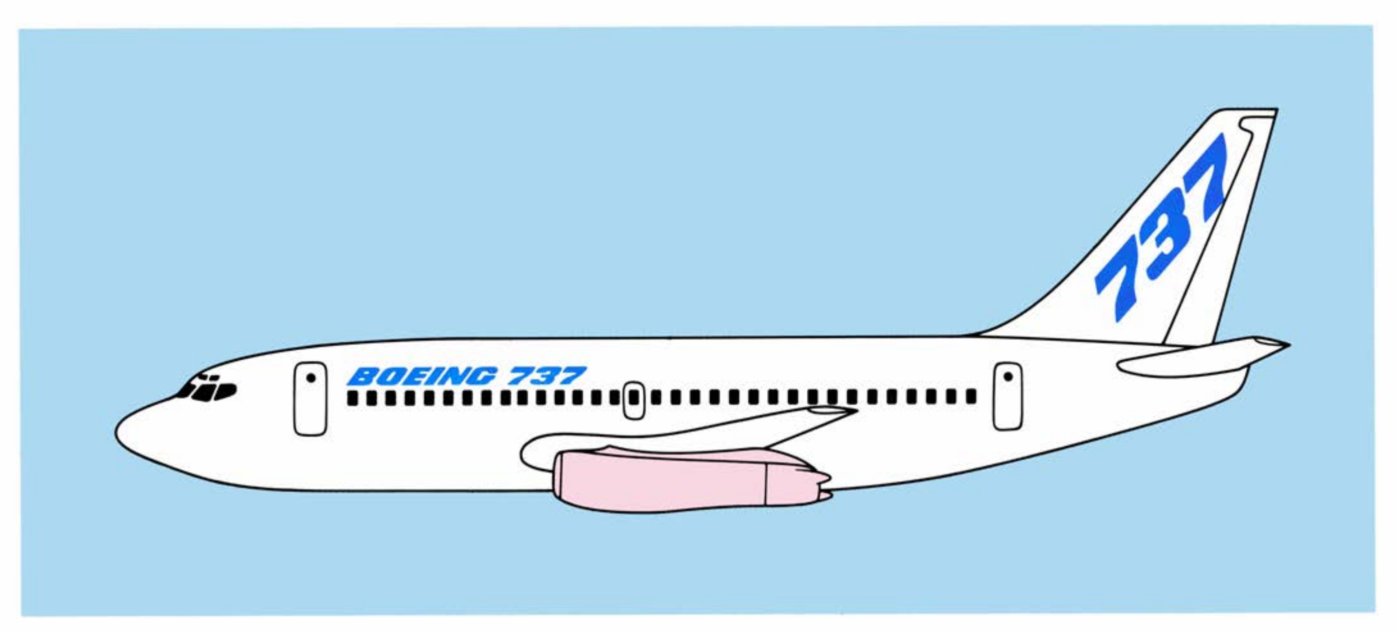
Según nuestro colega Paul E. Rubbert, de la compañía Boeing, la aplicación del cálculo por ordenador a la aerodinámica ha revolucionado el proceso de diseño, hasta el punto de que el superordenador ha adquirido en la industria aeronáutica una importancia comparable a la de los túneles de viento. Es más, la complejidad de la investigación informática que se lleva a cabo en el mundo entero en el campo de la aerodinámica sitúa actualmente a esa especialidad entre las de mayor ambición científica.

El superordenador ha permitido examinar detalles del flujo aerodinámico imposibles de medir en un túnel de viento tradicional o calcular con ordenadores comunes. Además, si se introduce el complejo volumen de datos en un superordenador, se obtiene una representación gráfica en tres dimensiones que, visualizada en una estación de trabajo adecuada, ofrece una amplia gama de ángulos y de perspectivas. Por esta razón en la Boeing se llevan a cabo de forma rutinaria análisis asistidos por ordenador de las distintas secciones de los aviones: fuselaje, ala, el montante que sostiene el motor y la barquilla de éste. Los ensayos en el túnel de viento indican la presencia de arrastres, pero el ordenador distingue las diversas componentes de esos arrastres y ofrece información precisa sobre sus contribuciones relativas. Por primera vez, el análisis por ordenador brinda informaciones que escapan a la experimentación física.

Los superordenadores desempeñaron un papel de primer orden en el diseño del Boeing 737-300, que entró en servicio en 1984, 17 años después de que se comercializase su predecesor, el Boeing 737-200, que tan buenos resultado ha dado. Ante la demanda de aviones de eficacia superior, en lo que al consumo de combustible se refiere, Boeing optó por una simple extensión del diseño básico del Boeing 737-200 y la sustitución de sus motores por otros más eficaces. Pero surgió un problema técnico: los nuevos motores eran mucho mayores que los originales.

Era evidente que no podían montarse los nuevos motores en las antiguas sujeciones de las alas del 737-200. Si así se hubiera hecho, para asegurar el espacio libre adecuado habría tenido que alargarse el tren de aterrizaje, añadiéndose un peso excesivo, obligando a realizar demasiadas modificaciomes estructurales e incrementándose considerablemente el costo del aparato, solución a todas luces inadecuada. Los ingenieros razonaron que, si se montaban los motores por delante de

1. CIRCULACION DEL AIRE SOBRE LA SUPERFICIE de un avión; puede modelarse en un superordenador y representarse en una imagen tridimensional. Se simulan (*arriba*) las trayectorias seguidas por partículas que emergen del motor y del borde de salida del ala de un Boeing 737-300. Se determina el camino exacto que siguen las partículas disponiendo una superficie artificial plana detrás del ala y midiendo en ella las coordenadas de las partículas, operación que se repite en varios puntos del flujo de aire. El superordenador recrea también el flujo de aire global y las correspondientes presiones que se ejercen sobre el aparato (*abajo*). Aquí se muestra la presión del aire sobre un Boeing 767-200 durante el despegue y el aterrizaje, cuando los alerones interiores ("flaps") están extendidos. Las zonas donde la presión del aire es menor (y el flujo mayor) aparecen en rojo; las áreas de máxima presión (y menor flujo de aire) se indican en púrpura.



2. LOS MOTORES DE PROPULSION del Boeing 737-200 se insieren en barquillas de pequeño diámetro montadas debajo de las alas. Estos motores consumen más combustible que los propulsores que se alojan en barquillas de diámetros mayores, porque su rozamiento con el aire es relativamente grande.

las alas, en vez de debajo, se obtendría suficiente holgura, sin incrementarse el peso y el costo. Sin embargo, el plan debía superar un obstáculo importante: las pruebas realizadas en el túnel de viento habían mostrado repetidamente que al montarse los motores cerca del borde de ataque del ala el arrastre aumentaba hasta niveles inaceptables. Rubbert afirma que, durante 20 años, los especialistas en aerodinámica han ensayado ese tipo de proyectos modificando el diseño de la barquilla de los motores en el túnel de viento, pero siempre habían fracasado.

Como era de esperar, el análisis por ordenador señaló que el acoplamiento del motor al ala creaba un arrastre excesivo. Sin embargo, a diferencia del túnel de viento, el análisis reveló el mecanismo de flujo real que generaba el arrastre, secreto que había quedado oculto durante dos décadas de ensayos en el túnel. Manipulando variables en el ordenador, los ingenieros sopesaron distintos proyectos con un nivel de detalle muy superior al que hubiera permitido el túnel de viento. A partir de esos estudios computacionales dieron una solución innovadora al emplazamiento del motor y al diseño de su barquilla. Modificaron el proyecto modelando cuidadosamente la parte superior de éste: le dotaron de una configuración asimétrica, en forma de pera, que, según se demostró, eliminaba casi por completo la interferencia con el flujo normal de aire por encima del ala. Se obtuvo así un motor acoplado al ala cuyo arrastre se había reducido de modo significativo.

El desarrollo del Boeing 757, cuya primera entrega data de 1982, ofrece otro buen ejemplo de la potencia que ha convertido a los superordenadores en importantes herramientas del diseño de aviones. Se había programado su aparición para seis meses después que la del modelo 767, por lo que la compañía decidió incorporar la misma cabina en ambos. La razón de ello era puramente pragmática: se exige a los pilotos que realicen un cursillo específico para cada tipo de cabina (y para cada tipo de panel de instrumentos) donde vuelan, lo que obliga a preparar un programa de entrenamiento distinto para cada modelo de avión. Si salían ambos modelos con la misma cabina, a los pilotos les bastaría un solo cursillo para familiarizarse con los dos aviones, con el consiguiente ahorro en dinero y tiempo. Sin embargo, a la Boeing se le planteaba un problema peliagudo: el 757 era un avión más estrecho y pequeño que el 767. ¿Cómo podía adaptarse al 757 la arquitectura de la moderna cabina diseñada para el 767, dados los distintos tamaños de ambos aparatos?

El flujo de aire alrededor de la zona donde se unen la cabina y el cuerpo del avión determina el rendimiento del aparato; un exceso de arrastre en esa zona incrementa de forma significativa el consumo de combustible y el ruido que llega al interior. Resultaba, pues, esencial el análisis aerodinámico detallado del 757 (con una nueva cabina que alojara los componentes más importantes de la del 767) para determinar la viabilidad del prototipo y, en caso afirmativo, cuál debía ser el proceso a seguir.

El procedimiento habitual para encontrar la mejor solución (es decir, la más aerodinámica) del problema de diseño consistía en construir varias maquetas de las secciones, con distintos puntos de unión, y someterlas a prueba en el túnel de viento; sin embargo, los estrictos plazos de producción del aparato dejaban poco tiempo. Por esa razón, las opciones de diseño se analizaron en un superordenador. En cuestión de pocos días (en contraste con el período de varios meses que hubiera exigido el estudio en el túnel de viento) se encontró un diseño de la cabina que alojaba la del 767 en la estructura, más estrecha, del 757 y satisfacía los criterios de eficacia referentes al tamaño y al flujo de aire. Los ingenieros confiaban tanto en los resultados obtenidos que recomendaron la fabricación del nuevo avión antes de someterlo a verificación experimental en el túnel.

El último avión de la Boeing, el 7J7, aún en fase de desarrollo, abre una nueva era en la propulsión a chorro. Los motores del aparato se han situado junto a la cola e impulsan dos conjuntos iguales de paletas curvas fijadas a la parte posterior de la barquilla de los motores. Las paletas externas resultan más eficaces que las que equipan comúnmente a los aviones, encerradas en los motores: permiten un considerable ahorro de combustible sin mermar las

prestaciones de los aeroplanos actuales. Tan fundamental cambio del diseño del nuevo modelo no hubiera resultado viable sin el concurso de los superordenadores. El análisis estructural de ciertos componentes, el análisis electromagnético y el diseño de sistemas de control constituyen otras tantas facetas del diseño aeronáutico que se sirven de un complejo modelado matemático y del superordenador.

La inmensa mayoría de las empresas automovilísticas se han decidido por el empleo de superordenadores en el diseño aerodinámico de sus vehículos. La compañía Ford hizo uso exhaustivo de esas máquinas al proyectar su modelo Taurus de 1986. Además del ahorro de combustible que beneficiará al usuario, se estima que la propia empresa ha ahorrado varios millones de dólares en costes de diseño, sobre todo por haber prescindido de la construcción de múltiples prototipos. Los fabricantes de automóviles se sirven también de superordenadores en la simulación de choques. Modelando en pantalla los daños de las colisiones, ahorran la gran cantidad de dinero que cuestan los valiosos prototipos que habría que destruir.

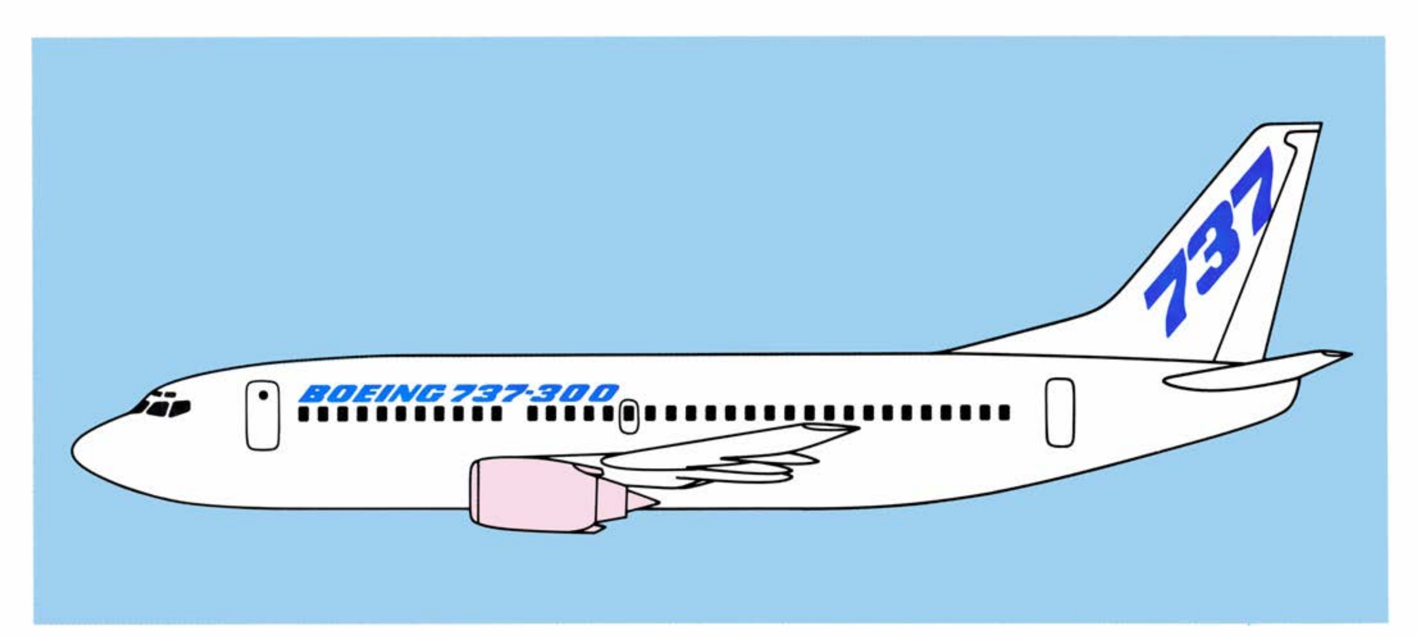
Probablemente sea el sector petroquímico el principal ejemplo de utilización industrial de los superordenadores. Desde finales de la década de 1970 se recurre a ellos en la prospección y extracción de petróleo. A partir de registros sísmicos almacenados en soporte magnético se van descubriendo los rasgos geológicos asociados con bolsas de petróleo, lo que permite a los ingenieros identificar las zonas de perforación con un grado de exactitud mucho mayor que el de los métodos tradicionales, incrementándose, por consiguiente, la probabilidad de descubrir un gran yacimiento. El costo de la perforación es tan alto (superior a 15 millones de dólares por pozo continental y notablemente superior en pozos marinos) que las cargas que comporta el ordenador resultan relativamente insignificantes al lado de los beneficios que aporta. También desempeñan los superordenadores un importante papel en la extracción del crudo; incrementan el ritmo de extracción modelando la dinámica de fluidos del flujo y la separación química.

Los superordenadores, muy dúctiles, pueden aplicarse en general a cualquier problema donde el uso de modelos matemáticos complejos aporte información nueva sobre el diseño de un producto o la mejora de su rendimiento. Se han empleado en el modelado de estructuras moleculares e imágenes de radar, así como para crear imágenes cinematográficas (por ejemplo, las escenas de la película *The Last Star Fighter* se generaron enteramente por superordenador). También se han empleado en la simulación de vuelo y en programas de entrenamiento de pilotos y, más recientemente, los han adoptado grandes compañías financieras para analizar sus carteras de inversiones.

En el campo de la arquitectura, los superordenadores modelan las tensiones y la elasticidad frente a los terremotos de grandes proyectos ingenieriles, como puentes, muros de contención de reactores nucleares, grandes polideportivos y rascacielos. También empiezan a aplicarse en la predicción meteorológica: se combinan los pormenores de las condiciones climáticas y geográficas con información real de tormentas y, una vez introducidos en el ordenador, se obtiene una representación tridimensional del movimiento de las borrascas. Los meteorólogos predicen así la trayectoria que probablemente seguirá la tormenta y avisan con mayor precisión qué áreas de población se verán afectadas.

Hemos descrito algunos de los logros asociados a la aplicación de los superordenadores en la industria. Esos éxitos no se han alcanzado sin esfuerzo, ni son tan amplios como podría haberse esperado. Para entender las razones que nos llevan a creerlo así convendrá repasar las limitaciones de los superordenadores.

El término superordenador se usa con dos sentidos. Define el primero aquellos ordenadores que tienen la mayor potencia en bruto, medida en términos de rapidez y de tamaño de memoria; en su segunda acepción se destaca la arquitectura interna del mismo, es decir, los componentes físicos que lo forman. En este artículo utilizamos el término superordenador refiriéndonos a la primera categoría: máquinas do-



3. EL DESARROLLO de motores mayores y más eficaces obligó a rediseñar el 737-200, resultado de lo cual fue el 737-300 que se muestra aquí; los motores se montan delante del ala, una posición que el análisis tradicional en el túnel de viento obligaba a rechazar, porque generaba mucho arrastre. Con ayuda del superordenador, los técnicos de Boeing distinguieron las diversas fuerzas que actuaban sobre el motor y rediseñaron la barquilla para minimizar el arrastre.

tadas de una potencia máxima, que suelen también presentar una arquitectura avanzada. A las máquinas de la segunda categoría —ordenadores que, pese a estar construidos con una arquitectura más avanzada, son menos potentes— las denominaremos cuasi-superordenadores.

Los superordenadores se desarrollaron para resolver complejos problemas numéricos. Por ello, la velocidad a la que operan constituye un parámetro fundamental de su potencia. Uno de los primeros prototipos de superordenador, construido a comienzos de la década de 1970, el modelo CDC 7600 de Control Data Corporation, trabajaba a cinco millones de operaciones de coma flotante por segundo, es decir, cinco megaflops (MFLOP). Los superordenadores actuales multiplican por 200 ese ritmo y alcanzan los mil millones de operaciones de coma flotante por segundo, es decir, un gigaflop (GFLOP). Los cuasi-superordenadores son aparatos considerablemente más lentos, pero su potencia, de entre 20 y 100 MFLOP, sigue impresionando.

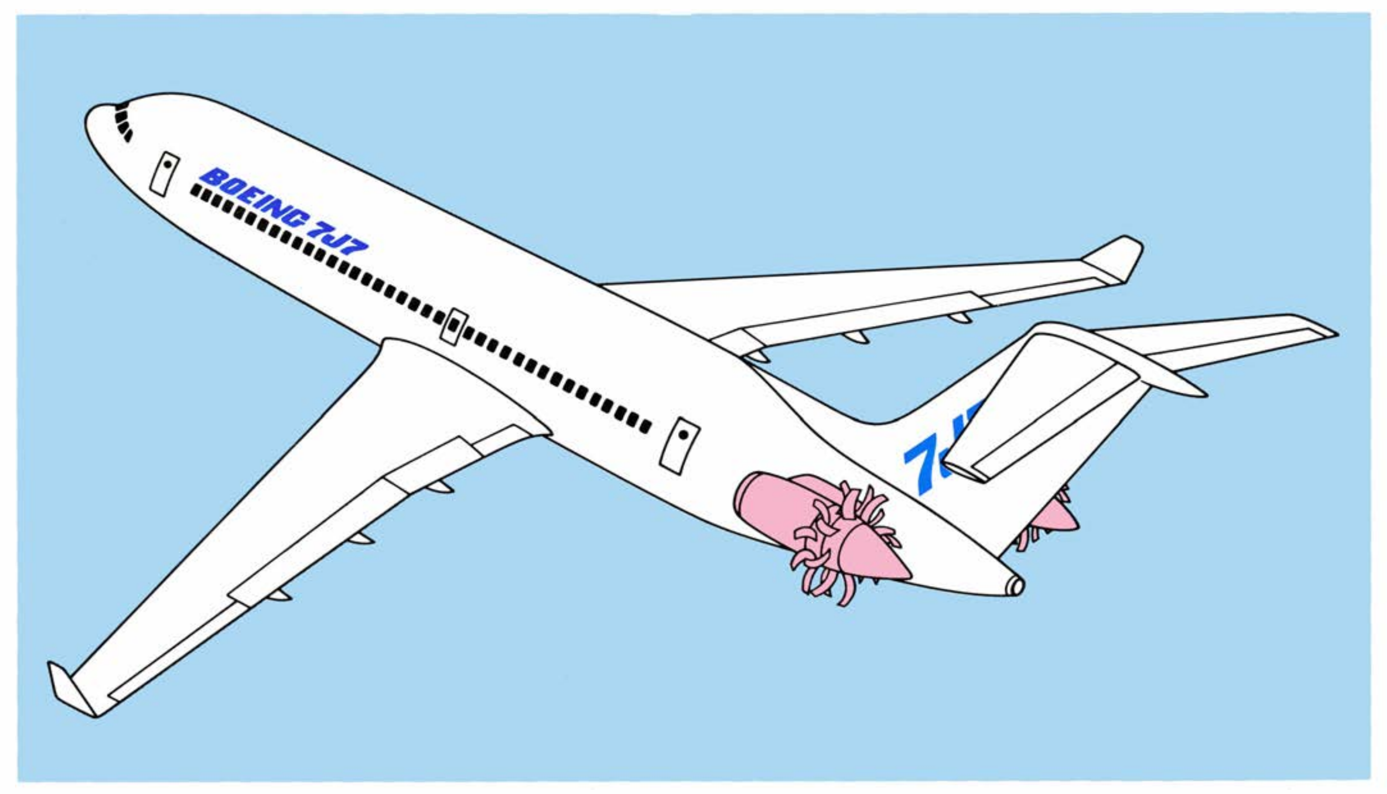
Un problema que ocuparía todo un año a un ordenador central convencional lo resuelve el superordenador en poco más de una hora. Tal abismal diferencia de la potencia de unos y otros sistemas responde a la arquitectura vectorial y en paralelo que configuran el soporte físico de la máquina. Las nociones de vector y arquitectura en paralelo son relativamente nuevos en el campo de las ciencias de la computación; se basan en dos aproximaciones: el procesamiento escalonado y el paralelismo (es decir, la reproducción de las unidades computacionales). [Para una exposición más detallada del tema véase "Arquitecturas avanzadas de computadores", por Geoffrey C. Fox y Paul C. Messina, en este mismo número.] Ambos enfoques proporcionan notables incrementos de la velocidad del ordenador, pero dan escaso resultado si no van acompañados del soporte lógico adecuado.

Un inconveniente de la arquitectura vectorial, que opera a modo de cinta transportadora, o de cadena de ensamblaje, es que la velocidad promedio a la que se ejecutan las operaciones depende del tiempo requerido para completar el primer cálculo que entra en el sistema y del número de cálculos que han de procesarse en la línea de transmisión. Se trata de un punto sutil, aunque importante. Los programas de ordenador que minimizan el número total de cálculos necesarios para un problema analítico, pero han de realizar un largo flujo de cálculos (denominados vectores largos), defraudan en ocasiones las previsones.

En teoría, y en lo que se refiere a la velocidad, el potencial de los procesadores en paralelo, en los que varios ordenadores interconectados atacan simultáneamente un problema, es ilimitado. Es decir, *n* ordenadores resolverán un problema en la *n*ésima parte de tiempo que emplearía un solo computador en resolverlo. Pero, al igual que ocurre con "el principio del millón de monos" en el área de gestión empresarial (un proyecto que en circunstancias normales ocuparía a una persona 10 años no pueden realizarlo un millón de personas en un minuto), el soporte físico paralelo no ha satisfecho todas las expectativas.

La arquitectura en paralelo y la vectorial se introdujeron con el ánimo de dar un gran salto en la potencia del cómputo. La evaluación de la capacidad de un superordenador en términos exclusivos del número de operaciones por segundo que es capaz de realizar equivale a ponderar la calidad de varios coches deportivos ciñéndose a su velocidad máxima, sin tener en cuenta la aceleración, el radio de giro máximo y las condiciones bajo las cuales pueden utilizarse, entre otros factores.

La potencia de un ordenador se basa en distintos considerandos: tamaño de



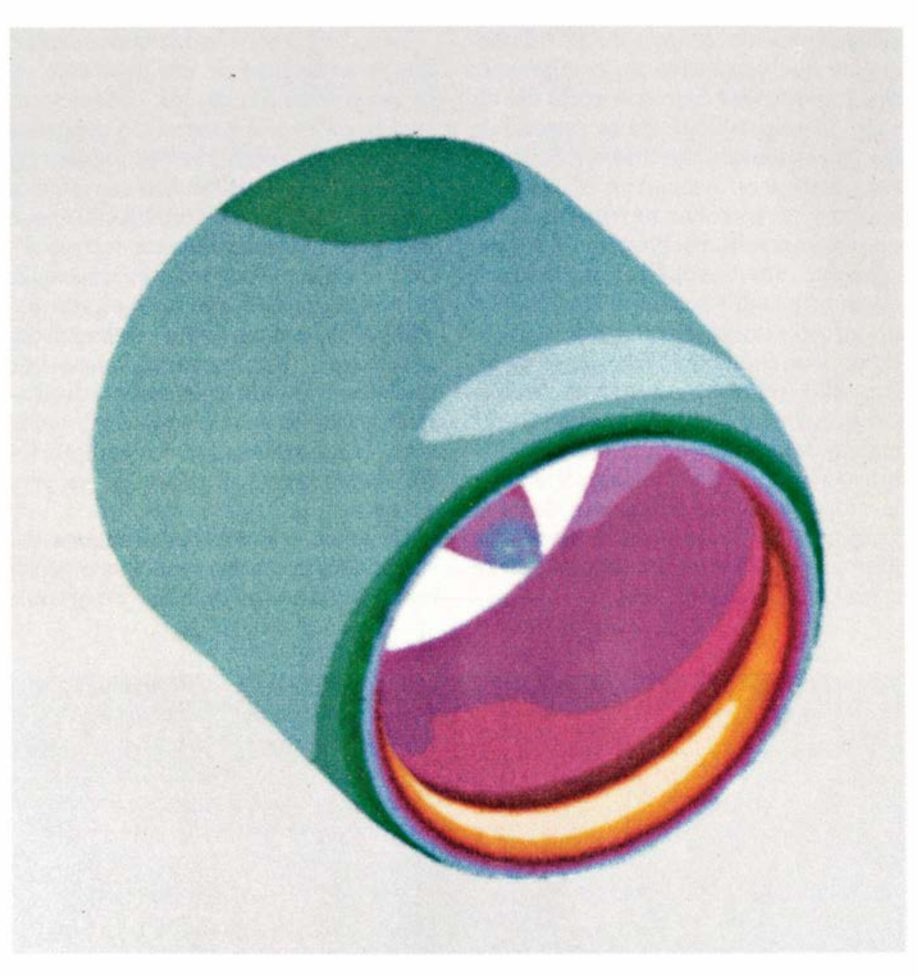
4. DIBUJO ESQUEMATICO de un nuevo avión, el Boeing 7J7 aún en diseño, que representa una nueva era en la propulsión a chorro. Los motores del aparato, montados junto a la cola, poseen dos conjuntos de paletas, que son más eficaces que las que alojan los motores actuales; permiten ahorros significativos de combustible sin mermar las prestaciones del aparato. El diseño no hubiera sido posible sin el análisis de muchas variables realizado en un superordenador.

la memoria, facilidad para manejar pequeños conjuntos de números (denominado el tamaño del vector pequeño), velocidad a la que pueden procesarse cantidades discretas de datos (denominada potencia escalar de la máquina) y capacidades del soporte lógico con que es compatible. Los computadores que trabajan bien en los peldaños más altos de esa escala (procesan grandes cantidades de datos con rápidez) no siempre lo hacen adecuadamente en los más bajos de la misma, cuando han de procesar pequeñas cantidades de datos. Por desgracia, la mayoría de las aplicaciones requieren del ordenador ambas facultades.

Así pues, la velocidad no es el único factor que determina el rendimiento de un ordenador. Depende también de otros dos: el tamaño de la memoria y la velocidad con que los datos fluyen de la memoria hacia las unidades internas de cálculo.

La memoria es un factor determinante a la hora de evaluar el rendimiento de un ordenador; la velocidad a la que se realizan los cálculos frecuentemente depende del tamaño de ciertos bancos de memoria de la máquina. Los cuellos de botella suelen deberse al propio volumen de datos en manipulación. En realidad, el movimiento de grandes cantidades de datos desde un almacenamiento secundario hacia la memoria, y viceversa, puede convertirse en el proceso que consuma más tiempo en cualquier problema de computación. Pese a que el tamaño de la memoria de los superordenadores crece con un ritmo comparable al de la potencia de cálculo, en algunos casos resulta todavía insuficiente. El modelo CDC 7600 citado antes tenía una capacidad de memoria de 500.000 números de 15 dígitos. El Cray-2, construido casi 15 años más tarde, presenta una capacidad 512 veces mayor: almacena en su memoria 256 millones de números de 15 dígitos. Pero al incrementarse la complejidad de los cálculos, memorias de varios miles de millones de palabras pueden resultar insuficientes para las necesidades futuras. Trabajar con semejantes tamaños exige cambios radicales del diseño de los programas.

Importa especialmente acoplar la capacidad de un superordenador a las necesidades de su usuario. Por ejemplo, al especialista en aerodinámica interesado en la medida de las tensiones de un ala no le preocupa la velocidad máxima de la máquina, sino la velocidad con que puede crear un complicado modelo matemático del ala y someterlo a una serie de pruebas de tensión. Por consiguiente, la rapidez con que el ordenador resuelve un problema complejo en el que intervienen muchas variables puede importar más que la velocidad con que procesa un conjunto ordenado de números. Volviendo a la comparación con el coche deportivo, no es lo mismo evaluar su rendimiento en una sinuosa carretera de montaña que en una autopista recta y plana.



5. BARQUILLA AISLADA, modelada en tres dimensiones por medio de un superordenador, que permite la cómoda modificación de variables. Los colores corresponen a distintos valores de la velocidad del flujo de aire dentro y fuera de la barquilla durante el vuelo. La velocidad del aire es máxima en el extremo interior frontal de la pieza (*amarillo*), donde el aire se canaliza hacia el revestimiento y hacia el interior del motor del avión. Es mínima hacia el extremo posterior de la barquilla (*púrpura*) e intermedio en el exterior (*verde*).

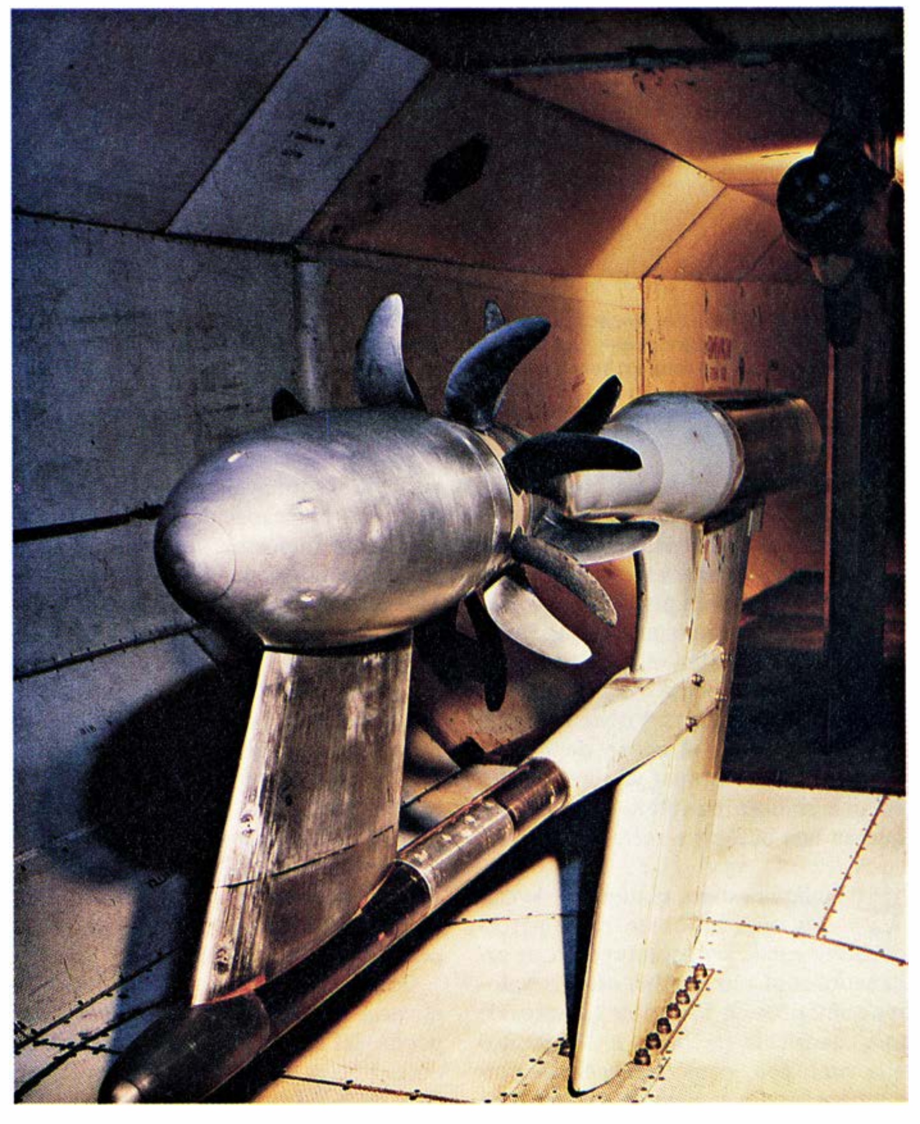
En cualquier caso, el número de cálculos producidos en un superordenador puede ser asombroso. Los ordenadores al uso pueden ofrecer valores numéricos de 50 puntos distintos de una cuadrícula (medidos en 20 puntos a la vez), pero empequeñecen al compararlos con los superordenadores, que generan datos de 50.000 puntos de una cuadrícula midiendo 200 por vez. Por fortuna, con los superordenadores, los fabricantes han desarrollado avanzadas terminales de trabajo capaces de convertir la información en imágenes gráficas; sin esa capacidad resultaría virtualmente imposible interpretar la ingente cantidad de datos producidos. Más aún, unen en secuencia varios diseños gráficos y los proyectan a modo de película que muestra los datos cual si se tratara de una simulación realista de un proceso físico. Por ejemplo, el flujo de partículas de aire que rodean un ala aparece como un flujo continuo, lo que facilita la visualización de las fuerzas que actúan sobre ella. Y lo que importa tanto o más: puede modificarse el diseño manipulando las variables y haciendo aflorar las componentes específicas de la turbulencia.

Como en todos los ordenadores, la eficacia de los superordenadores depende del soporte lógico de que se les dote. Ello es especialmente cierto en el área de la fabricación. Modificar o rehacer los programas empleados en investigación básica con superordenador resulta una tarea bastante fácil; cuesta mucho más modificar un programa de 100.000 líneas sometido a estrictos pro-

cedimientos de control de la fabricación y que constituya un componente fundamental del ciclo industrial del diseño de un producto. En la mayoría de las aplicaciones industriales, el tiempo necesario para evaluar un nuevo programa y para probar su fiabilidad y su precisión resulta prohibitivo. Por consiguiente, interesa que el soporte lógico de un ordenador posibilite la transición de un tipo de arquitectura de computador a otro. Sin embargo, en la práctica ello no siempre resulta viable. Consecuentemente, de la misma forma que ha habido que demoler algunas fábricas para acoger la automatización y la modernización, habrá que derribar programas anticuados, para dejar sitio libre a una nueva generación adaptada a los nuevos ordenadores.

Se ha intentado diseñar herramientas que se encarguen de esa transición de la programación de los ordenadores tradicionales a la generación avanzada de máquinas, pero el éxito logrado ha sido limitado. Una de esas herramientas es el "compilador inteligente", que vectoriza automáticamente instrucciones, es decir, transforma instrucciones de un lenguaje de alto nivel en instrucciones específicas para ordenadores vectoriales. Sin embargo, más que identificar nuevos algoritmos o diseñar nuevas formas de resolver un problema con el superordenador, los compiladores redistribuyen el orden de un programa.

La única solución del problema del soporte lógico, opinan muchos especialistas, consiste en redactar programas en un nuevo lenguaje específicamente diseñado para máquinas vectoriales o de proceso en paralelo. Habrá que remplazar los lenguajes tradicionales, como el FORTRAN, principal soporte de la programación científica durante los últimos 20 años, por un lenguaje más adecuado a la avanzada tecnología de los superordenadores. Pero el intervalo que separa el desarrollo de un nuevo lenguaje y su aplicación por parte de los fabricantes y de los usuarios que, en última instancia, deban reescribir, verificar y validar miles de millones de dólares de programas, resulta enorme, si no absolutamente impracticable. La lentitud acompaña incluso a los cambios menores de lenguaje. Por ejemplo, la nueva versión del FORTRAN, la 8X; apareció a finales de la década de 1970, pero es poco probable que se acepte y se generalice su empleo antes de principios de los 90. Ni siquiera entonces el FORTRAN 8X pasará de ser un lenguaje limitado, en lo que atañe a su capacidad de operar con un rendimiento máximo en ordenadores de arquitectura avanzada. Por tanto, la evolución de nuevos lenguajes de programación sólo resolverá, y lentamente, una pequeñísima parte del problema.



6. MODELADO DE COMPONENTES AERODINAMICOS en el túnel de viento. Aunque se complementa cada vez más con modelos de ordenador, aún desempeña un importante papel en los estadios finales del diseño. El motor que se muestra porta dos juegos de paletas que giran en sentidos contrarios y se ha instalado en un montante donde puede someterse a prueba a velocidades que superan los 900 kilómetros por hora. Durante la prueba, se introducen en el túnel unos sensores que miden la velocidad del aire por encima de la barquilla; los que salen del techo portan micrófonos que registran el ruido generado por las paletas.

Otra alternativa corresponde a la categoría del soporte lógico que comprende programas formados por bloques de cálculo. Esos "sillares de cálculo" son algoritmos pre-programados que una determinada comunidad de usuarios acepta como programas convencionales. Los programas "prefabricados" ofrecen tres ventajas: utilizan un lenguaje de interfase convencional (y por consiguiente pueden ejecutarse en cualquier máquina); están cuidadosamente adaptados a la arquitectura del soporte físico y se han diseñado para resolver los cálculos habituales, así operaciones con matrices, transformaciones de Fourier y operaciones de clasificación —tareas que constituyen el esqueleto de la mayoría de los problemas analíticos complejos. Puesto que buena parte del tiempo de cómputo se dedica a ese tipo de cálculos, los bloques prefabricados (de construcción relativamente barata) ahorran enormes cantidades de tiempo y dinero. El superordenador Cray X-MP de la Boeing, y sus otros ordenadores industriales avanzados, utilizan anualmente centenares de millones de veces las dos "bibliotecas" de bloques prefabricados de la empresa, la matemático/estadística, BCSLIB, y la biblioteca de subrutinas vectoriales.

Las ventajas de los bloques prefabricados no acaban en su capacidad de manejar cálculos complejos: permiten verificar la fiabilidad de una ruta particular de diseño antes de desarrollar el programa completo necesario para su aplicación. Comparando distintas opciones del problema antes de su ejecución, el usuario ahorra no poco tiempo y dinero.

Frente a la potencia y el enorme crecimiento de los superordenadores, otro obstáculo entorpece su adopción generalizada: la dificultad de integrar sus tres componentes, a saber, el usuario, la terminal de trabajo gráfica y el propio superordenador. Los ordenadores tradicionales suelen conectarse con sus usuarios a través de líneas telefónicas que intercambian los datos a una velocidad de 1200 baudios, es decir, bits por segundo. Pero comunicar un superordenador, capaz de manejar enormes volúmenes de datos, a 1200 baudios es como pretender vaciar el océano con una cañita; por cable no pueden transmitirse a grandes distancias elevadas cantidades de datos con suficiente rapidez.

Apreciar los límites que ello impone resulta de suma importancia para la supercomputación: entrados los datos que ocuparán sólo cinco minutos de proceso en el superordenador (y que entretendrían doce horas a un ordenador común) quizá se produzca un retraso de doce horas en la obtención de los resultados. En la actualidad, las conexiones de larga distancia están capacitadas para procesar 56 kilobits de información por segundo; sin embargo, incluso a esa velocidad resultan demasiado lentas para los superordenadores. La creación de secuencias en movimiento (que muestren, por ejemplo, el flujo de partículas alrededor de un ala) puede consumir más tiempo: en vez de analizar una sola vez un grupo único de variables, el superordenador debe resolver una serie de cálculos y luego generar imágenes —que se suman al flujo de datos de la red. Por esa razón, algunos usuarios han decidido prescindir de la comunicación telefónica y optan por la vía más rápida de enviar y recibir datos mediante un servicio de mensajería. Si bien no constituye la solución ideal, se trata de una posibilidad que cuenta cada vez con más adeptos.

Los Alabama Supercomputer Network Services, de Huntsville, cuya puesta en marcha está prevista para enero de 1988, ejemplifican la adopción de una solución distinta. Se espera que su superordenador transmita información entre las distintas facultades de la Universidad de Alabama a una velocidad de 1,5 millones de bits por segundo. En una escala superior, continental, por ejemplo, un sistema de ese tipo resultaría económicamente prohibitivo, pero puesto que la red informática de Alabama consta de un número reducido de terminales se espera que el sistema funcione bien.

El centro Ames de investigación que la NASA tiene en Moffett Field, California, donde se encuentra el ordenador más potente del mundo, ha optado por dar una solución distinta al problema de las comunicaciones a larga distancia; solución que se adecúa a algunas de sus investigaciones. Se han colocado potentes terminales de trabajo en las proximidades del superordenador central, obviándose la necesidad de establecer líneas de larga distancia. En la sede central de la Boeing Computer Services Company, de Bellevue, Washington, se ha adoptado una solución similar. Un grupo de especialistas en aerodinámica está conectado con el superordenador Cray X-MP mediante una serie de terminales gráficas de trabajo alojadas en un laboratorio cercano al superordenador. Tales disposiciones han reducido a escasos hectómetros la distancia a la que ha de transmitirse la información.

En un segundo paso hacia la integración, el incremento de la productividad no se centra en los superordenadores, sino en las terminales de trabajo. Se trata de una disposición "cómoda", en la que el científico o ingeniero trabaja con un ordenador personal o una terminal de trabajo utilizando programas de aplicaciones comunes. Todas las comunicaciones con el superordenador las gestiona directamente la terminal de trabajo, empleando un soporte lógico que se comunica con el usuario en el lenguaje de su aplicación. La información se introduce directamente en el superordenador, donde se convierte en código vectorial; el usuario de la terminal no sufre las limitaciones del lenguaje de alto nivel. Que el análisis se realiza en un superordenador sólo se advierte por la complejidad de los resultados.

En la mayoría de las aplicaciones, la utilización de programas "cómodos" para el usuario topa con la lentitud de la red de comunicaciones y con la escasa potencia de los ordenadores personales. La Boeing ha desarrollado un servicio de ese tipo, The Pretoleum

Gallery Service, en un paquete integrado de productos diseñados para la industria del petróleo. The Petroleum Gallery Service ofrece una única fuente de datos, aplicaciones y potencia de cálculo para los geólogos que cartografían y analizan los yacimientos.

Pese a lo limitado de esos avances, el verdadero potencial de la aplicación de superordenadores en la fabricación aflorará cuando se integre el análisis computacional, en el que tanto se han distinguido los superordenadores, con los procesos de producción, donde hasta ahora ha sido escaso su impacto directo. Se ve como algo apremiante la integración del análisis del diseño del producto con la propia producción.

Ciertas informaciones de importancia decisiva en el proceso de producción, como son las tolerancias del diseño, la optimización del costo de fabricación y la fiabilidad a largo plazo, raramente se consideran durante el análisis de un proyecto. Integrando las necesidades de la producción con el análisis inicial del diseño de un producto pueden alcanzarse compromisos que tengan en cuenta las necesidades de producción y el rendimiento del producto.

Por varias razones, tal integración no se ha producido aún. Los superordenadores son instrumentos jóvenes: han pasado del laboratorio a la industria apenas hace diez años y, por tanto, son demasiado recientes para que hayan influido sobre muchos de los productos que están saliendo de las cadenas de montaje. El proceso de diseño obliga a que la relación CAD/CAM permanezca estable durante el período de diseño entero, que a menudo es largo. Debe ejercerse a lo largo de muchos años un control estricto de la programación, componentes y datos. Los cambios rápidos de las arquitecturas de los ordenadores avanzados y las consiguientes modificaciones de la programación no encajan en esa exigencia de estabilidad. Productos como un avión comercial, por ejemplo, experimentan una transición muy larga desde su proyecto inicial hasta su fabricación final. Dada la interdependencia de sus componentes, resulta muy probable que la alteración del diseño de un componente obligue a cambiar otros. Por esa razón, varios estadios del proceso de diseño se "congelan" en ciertos puntos; no se admiten variaciones ulteriores y el proceso completo de CAD/CAM se "petrifica".

La situación podría cambiar cuando empiecen a producirse modelos más complejos, hoy aún en fase de desarrollo. Pero debe mencionarse otro motivo de preocupación: a pesar de los enormes beneficios de la integración CAD/CAM, los dos componentes no siempre resultan mutuamente compatibles. Por ejemplo, el modelado aerodinámico quizá dicte que, para que la eficacia sea máxima, un ala deba tener cierta forma, pero ese diseño a lo mejor entra en contradicción con otras exigencias de la producción: puede que el metal no se curve según las previsiones del superordenador, o quizá sea demasiado pesado o demasiado quebradizo para el proyecto.

En último lugar, los modelos computacionales de análisis del flujo aerodinámico ponen a prueba incluso la capacidad de los superordenadores actuales; procesar los cálculos extraordinariamente complejos con diversos parámetros de inseguridad y tolerancia, necesarios para la completa integración CAD/CAM, supera la capacidad de los ordenadores actuales. A la vista de esas limitaciones, no se han adaptado los superordenadores para ciertos problemas de cálculo, desaprovechándose así uno de los mejores atributos del superordenador: la velocidad con que manipula la información.

A la postre, los superordenadores quizá se integren con la computación propia de la inteligencia artificial, en la que no domina hoy el cálculo numérico, sino el simbólico; ésta se realiza en ordenadores especializados. En teoría, algún día podrá aprovecharse la inteligencia artificial para obtener mayor información sobre los enormes y complejos conjuntos de datos procesados por los superordenadores.

No erraremos al afirmar que la supercomputación se encuentra en un período de transición, y que ese cambio se registra simultáneamente en muchos aspectos: potencia, arquitectura, capacidades analíticas, competitividad e inteligencia artificial. Los superordenadores han revolucionado la industria aeronáutica hasta tal punto que sigue sin satisfacerse la demanda de ordenadores más rápidos y más potentes. Los próximos cinco años serán de particular importancia, y ello no sólo porque serán muchas las oportunidades para que se produzcan hallazgos tecnológicos revolucionarios, sino también porque en esta era de competencia internacional especialmente enconada la tecnología occidental quizá dependa de los superordenadores.

Juegos de ordenador

"After MAD*": juego computarizado de estrategia nuclear que acaba en un "dilema del prisionero"*

A. K. Dewdney

El lector dirige una superpotencia nuclear. El general jefe del Estado Mayor de las Fuerzas Armadas acaba de pasarle una hoja con un cuadro de números un tanto desconcertante. Muestra dos filas rotuladas "Nosotros" y dos columnas encabezadas "Ellos". El general se dirige a usted en estos términos:

"Esta matriz sintetiza en forma numérica nuestros análisis militares, económicos y políticos más perfectos. Como ve, hay en ella una casilla para cada posible combinación de acciones por parte propia y del enemigo. Por ejemplo, si nosotros atacamos y el enemigo no, la estimación prevista de ganancia por nuestra parte es de 63 puntos negativos ..."

El general duda un instante, pero al punto se anima.

"... mientras que la estimación de la ganancia enemiga se cifra en 74 puntos negativos."

Su jefatura de la nación le ha proporcionado a usted amplia experiencia en pronunciar discursos y en cortar cintas inaugurales. Aunque ha asistido a reuniones de planificación estratégica, más tarde ha confesado que las encontraba francamente complicadas. Y prosigue ahora el jefe del Estado Mayor:

"En breve, lo que tenemos aquí es una simple matriz de dos por dos que muestra las ganancias y pérdidas en el caso de que uno de los dos bandos, ninguno de ellos, o ambos lancen un ataque nuclear contra el otro."

En realidad, el lector se encuentra en la etapa inicial de "After MAD", un juego computarizado de estrategia nuclear inventado por Tad Homer-Dixon y Kevin Oliveau en 1985, en el Instituto de Tecnología de Massachusetts. La sigla MAD (además de expresar locura) corresponde a *mutually assured destruction*, es decir, destrucción mutua asegurada. "After" (después) alude a un punto del juego donde la destrucción ya no es tan mutua; al ir mejorando la precisión de las armas desplegadas, acaba resultando posible lanzar un primer ataque que deje al enemigo suficientemente malherido para que éste se vea imposibilitado de responder con otro contrataque nuclear importante. Así, crece en los jugadores la tentación de atacar sin haber sido provocados, hasta que se encuentran en la situación conocida por "dilema del prisionero" [véase "Temas metamágicos", por Douglas R. Hofstadter; INVESTIGACIÓN Y CIENCIA, julio de 1983]. Explicaré más adelante las peculiaridades de "After MAD".

Como ya he señalado, el elemento clave del juego es una matriz de dos por dos. El número de campos responde a que cada bando dispone de dos opciones, a saber, atacar y abstenerse de hacerlo. El contenido de la matriz cambia en cada etapa del juego, de tal modo que las cifras que figuran en ella resumen el equilibrio estratégico del momento. Cada entrada de la matriz consta de un par de números, uno para cada una de las posibles combinaciones de opciones elegidas por los dos bandos. El primer número de cada par es el que corresponde al bando propio en esa combinación de opciones; el segundo, el del bando contrario. Por ejemplo, si nuestro bando hubiera decidido abstenerse de atacar y el bando Ellos lo hiciera, la primera cifra sería inferior a la segunda.

Lógicamente, la matriz de juego subsiguiente es función de las combinaciones de jugadas decididas por una y otra parte en respuesta a la matriz actual. Cada matriz de juego conduce a otra, seleccionada entre cuatro en razón de las opciones que elijan los jugadores. Hay en "After MAD" 110 posibles matrices de juego o, si se quiere, 110 situaciones diferentes. Se distinguen cinco fases en el desarrollo de una partida:

Fase I (de destrucción mutua asegurada, DMA). Las armas estratégicas de cada bando sólo resultan eficaces contra objetivos "blandos", como centros industriales y de población.

Fase II (10 años posterior a la DMA inicial). Despliegue de misiles balísticos intercontinentales (ICBM) provistos de varios vehículos de reentrada a la atmósfera que pueden dirigirse independientemente hacia blancos distintos (MIRV). Siguen manteniéndose en reserva gran número de armas dirigidas contra objetivos blandos, pero cada bando ha adoptado lo que podríamos llamar "estrategia de contragolpe". Sus MIRV apuntan exclusivamente a los silos de misiles intercontinentales del enemigo.

Fase III (20 años después de la DMA inicial). La precisión de los MIRV ha mejorado enormemente. La noción de destrucción mutua asegurada ha perdido verosimilitud, pues quien decidiera atacar por sorpresa tendría garantizado poner fuera de combate el 75 por ciento del armamento estratégico de precisión del adversario.

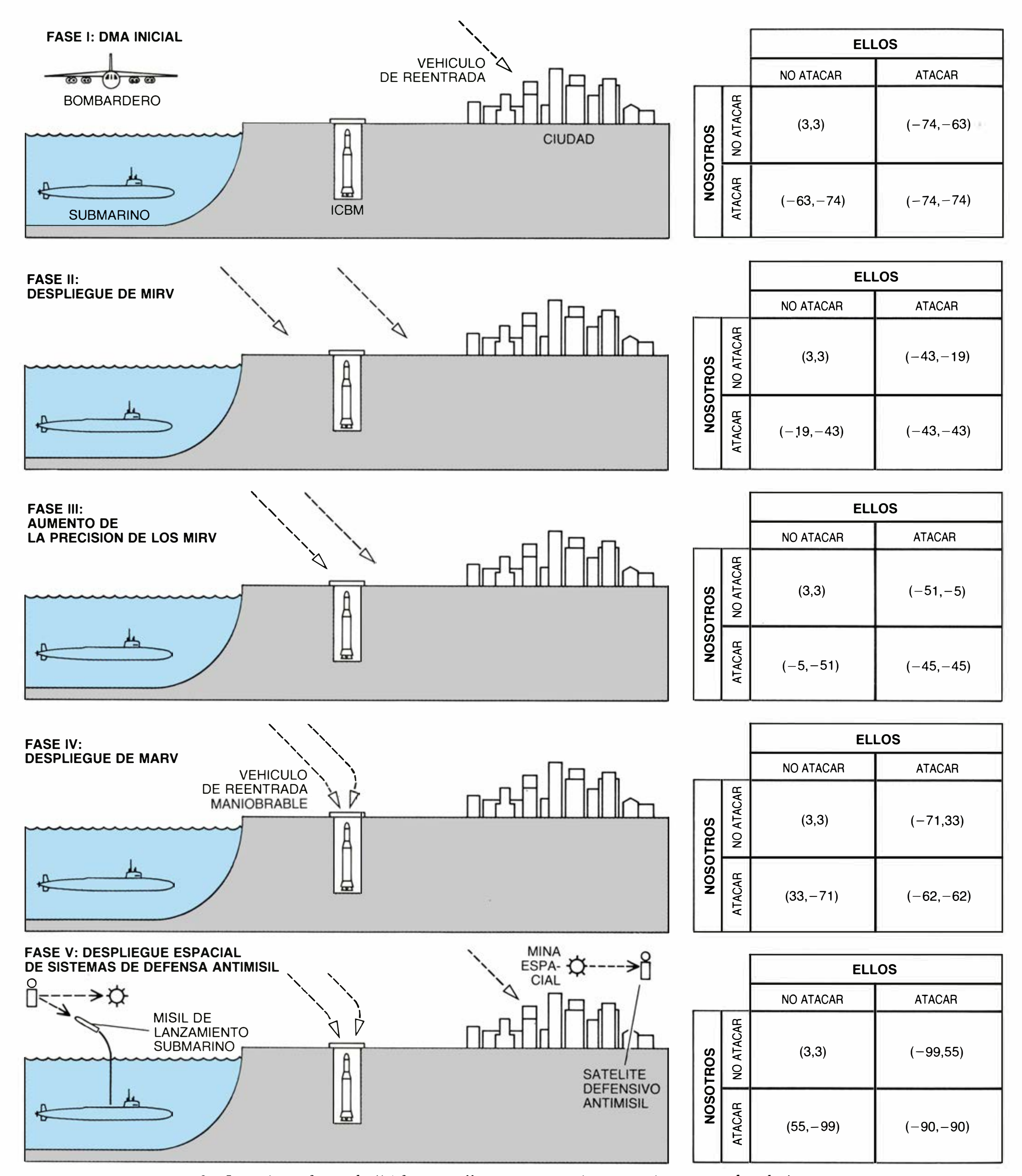
Fase IV (30 después de la DMA inicial). Ambos bandos han desplegado vehículos capaces de ser maniobrados tras su reentrada en la atmósfera (MARV) y de destruir prácticamente el 100 por ciento de los misiles intercontinentales del enemigo. La noción de destrucción mutua asegurada ha quedado obsoleta, pues ahora resultan militarmente atractivos los "ataques de contragolpe" asestados por sorpresa. La amenaza de represalias sobre los

		EL	
		COOPERAR	DEFRAUDAR
ELLA	COOPERAR	(3,3)	(0,5)
	DEFRAUDAR	(5,0)	(1,1)

1. Matriz del dilema del prisionero

núcleos de población ha perdido la capacidad de disuasión de esos ataque.

Fase V (45 años después de la DMA inicial). Los sistemas de defensa antimisiles balísticos desplegados en el espacio (SBMD) poseen capacidad para destruir por completo cualquier ataque de proyectiles intercontinentales que utilice como máximo el 50 por ciento de los cohetes de lanzamiento del adversario. Ambos bandos disponen igualmente de minas espaciales situadas en órbitas próximas a los componentes del sistema de defensa espacial del enemigo. Por tanto, el país que atacase primero podría destruir el sistema defensivo emplazado por el otro en el espacio, así como sus centros de población y misiles balísticos intercontinentales. El sistema defensivo espacial del atacante permanecería intacto y con capacidad para bloquear un posible contrataque lanzado por submarinos del enemigo.



2. Las cinco fases de "After MAD" y sus respectivas matrices centrales de juego

		ELLOS	
		NO ATACAR	ATACAR
NOSOTROS	NO ATACAR	(3,3)	(−71,33)
	ATACAR	(33,−71)	(−62,−62)

3. Matriz central de la Fase IV de "After MAD"

Cinco matrices de juego, a las que llamaremos matrices centrales, ponen de manifiesto las inestabilidades relativas de esas cinco fases [*véase la figura 2*]. Conforme se desarrolla el juego, si ninguno de los adversarios ataca, van pasando de una matriz central a la siguiente. No obstante, va creciendo progresivamente la tentación de atacar, pues aumentan los beneficios que se obtendrían de ello y se sabe, además, que el enemigo está experimentando la misma tentación. Las dos últimas matrices corresponden a sendos "dilemas del prisionero".

Como expuso Douglas R. Hofstadter en estas páginas, un "dilema del prisionero" comienza cuando, para interrogarlos, la policía detiene a dos sospechosos cerca del lugar donde se ha cometido un delito. Como es habitual, se les interroga por separado, para que no coordinen las declaraciones. Resulta que ambos son culpables, pero ninguno sabe si el otro confesará o no el delito. Si ambos se declaran culpables, las cosas se les pondrán feas a la pareja. En cambio, si sólo confiesa uno de ellos, quien suelte la lengua parecerá relativamente bondadoso y sufrirá una condena mucho más leve.

En lenguaje del dilema del prisionero diríamos que cada detenido tiene la opción de "cooperar" con su compinche, negando hallarse complicado en el delito, o, por el contrario, "defraudar" al otro, cantando de plano. El dilema del prisionero tiene asociada una matriz canónica [*véase la figura 1*]. He llamado a los detenidos El y Ella, convencido como estoy de que un toque de romanticismo hace más interesantes las discusiones técnicas. En el caso de que tanto Ella como El cooperen, ambos reciben en "pago" tres puntos. Si ambos defraudan, sólo reciben un punto cada uno mientras que, si solamente uno de ellos defrauda, quien lo haga recibirá cinco puntos, y el ingenuo que confió en el otro y cooperó, ninguno.

¿Cuál es la forma óptima de proceder en este juego? La evaluación de estrategias para el dilema del prisionero se efectúa realizando no una, sino muchas partidas. Al hacerlo así, el pago medio se convierte en cifra de mérito, o "calificación" de la estrategia.

En 1983, Robert Axelrod, del Instituto de Estudios de Política Pública de la Universidad de Michigan, celebró un campeonato de "dilema del prisionero". Los participantes en el torneo eran programas de ordenador que daban cuerpo a diversas estrategias de juego. Cada pareja de programas celebró 1000 partidas. El vencedor destacado del torneo fue un programa muy sencillo, bautizado OJO POR OJO. [Tras varios artículos, Axelrod publicó en 1984 sus resultados en *The Evolution of Cooperation*, libro del que existe traducción española: *La evolución de la cooperación*, Alianza Universidad; Madrid, 1986. N. del T.] La estrategia concretada en el programa consistía, sencillamente, en empezar cooperando y repetir después lo que el otro bando hubiera hecho en la jugada anterior.

Las matrices que figuran en "After MAD" son semejantes a las del dilema del prisionero. He sustituido las palabras "cooperar" y "defraudar" por las de "no atacar" y "atacar". Además, de ordinario los valores que figuran en las matrices son mayores, como ocurre en la matriz central de la fase IV de "After MAD" [*véase la figura 3*]. Las cifras de la matriz guardan entre sí la relación característica de los dilemas del prisionero: las cifras correspondientes a la "tentación" de defraudar (33), la "recompensa" (3), la "penalización" (−62) y la "cuota del ingenuo" (−71) forman una sucesión de números estrictamente decreciente. Se diría que los jugadores de "After MAD" habrían de aplicar la estrategia del "ojo por ojo", que no sólo aseguraría una puntuación razonable a cada jugador sino, por así decirlo, también la paz mundial.

Homer-Dixon y Oliveau se han tomado no pocas molestias para dar realismo al juego, tanto en lo que concierne a los valores asignados a las matrices como a la descripción textual de cada situación. En parte, desarrollaron el juego con el ánimo de que sirviera como instrumento de investigación con el que explorar los efectos que producirían los cambios tecnológicos sobre la relación estratégica entre la Unión Soviética y los Estados Unidos. Sin embargo, sus hipótesis iniciales referentes a cómo jugarían los estudiantes universitarios (entre otras personas) quedaron invalidadas por el desarrollo real de las partidas. Por lo general, para entrar en guerra los jugadores no esperaron a que las matrices llegaran a hacerse inestables: en el 35 por ciento de las 100 partidas jugadas en el Instituto de Tecnología de Massachusetts, uno de los bandos lanzó un ataque nuclear pese a no hallarse todavía en fase de destrucción mutua asegurada.

Es preciso hacer notar que los estudiantes que participaron en el juego habían sido previa y cuidadosamente instruidos, y se les exigía dar justificación escrita de sus jugadas. Por tanto, y en la medida en que la supervivencia del mundo propiamente dicho equivale a un único ensayo de lo mismo, tal vez hayamos tenido suerte. Describiré seguidamente de qué modo procedió Homer-Dixon para construir las matrices de juego, y cómo respondieron a ellas los jugadores.

Los pagos que figuran en cada matriz de juego de "After MAD" son resultado de tres componentes: la propiamente militar, el bienestar social y económico y el poder político. El mínimo de los pagos positivos, 3, está compuesto por una unidad positiva de cada componente. Tal pago supone que el país se encuentra en paz, con un pequeño cambio incremental de cada una de las tres componentes durante el breve período correspondiente a una jugada. El máximo número de unidades positivas o negativas que puede aportar una componente cualquiera es 33.

He aquí cómo calculó Homer-Dixon el pago de la "tentación" que figura en la matriz central de la Fase III del juego, que representa la puntuación recibida por el bando que opta por atacar mientras el otro se abstiene de hacerlo. La ventaja militar se estimó inicialmente en +5, pero debió disminuirse en 3 unidades en razón del costo de las posibles represalias. El valor 3 se obtuvo como sigue. Si un bando lanzase un ataque total de contragolpe podría llegar a destruir el 75 por ciento del material estratégico del otro. En tal caso, el que ataca primero podría inflingir a su adversario una pérdida de $75/100 \times 33 = 24$ puntos (aproximadamente). Puesto que, tras haber sufrido el ataque inicial, al bando atacado sólo le queda alrededor del 25 por ciento de su armamento estratégico, las pérdidas que podrá inflingir a su atacante serán sólo de $25/100 \times 33 = 6$ puntos. A esa pérdida se le aplica un factor de descuento de dos puntos, pues el país atacado podría responder hasta con el 25 por ciento de sus fuerzas estratégicas.

La componente militar, +2, que subsiste al restar 3 (el costo de la potencial represalia) de los 5 puntos de ganancia militar, se suma a las aportaciones de las otras dos componentes: −2 para el bienestar social y económico y −5 para la potencia política. Es de presumir que el país atacante perderá también algo de industria y de población tras la respuesta del enemigo. Su gobierno sufriría (en un mundo cuerdo) la repulsa unánime de los demás, por haber sido el primero en atacar. En suma, el atacante ha ganado $+2 - 2 - 5 = -5$ puntos.

Los escenarios, o teatros de operaciones, utilizados en "After MAD" enuncian cuidadosamente las consecuencias de carácter estratégico que comporta la situación vigente. Homer-Dixon y Oliveau explican igualmente la forma en que los valores de la matriz reflejan tales implicaciones. Dado que hay 110 matrices, habrá también 110 escenarios, que se comentan a los jugadores a su llegada a la correspondiente matriz.

Se desencadena la guerra cuando uno de los jugadores (o ambos) deja de cooperar y decide atacar al otro. Alcanzado ese punto, el juego pasa de la matriz central asociada a tal acción a una de las matrices de ramificación. La guerra puede terminar en destrucción de ambos bandos (con lo que concluye el juego) o haciendo retroceder a los adversarios hasta una matriz central anterior, donde tratarán de recomponer los pedazos de sus destrozadas existencias.

Aunque se instruyó a los jugadores para que procuraran en lo posible maximizar su puntuación, éstos tomaron en cuenta, según parece, otras consideraciones. Incapacitados para verse o comunicarse, su conducta mostró a veces rasgos paranoides. Veamos, por ejemplo, cómo sobrevivieron dos jugadores, a quienes llamaremos sencillamente Fila y Columna, tras una partida de "After MAD".

La actuación de Fila fue de pacifista a ultranza. Como indicaban los registros de su ordenador, Fila no estaba esforzándose por maximizar su tanteo, sino por atenerse a principios morales. "Mi juego no pretenderá sumar puntos. Me propongo, en cambio, jugar casi siempre de acuerdo con mis principios. Lo más probable es que casi siempre coopere... Espero que el otro jugador actúe como yo."

El registro correspondiente a las primeras jugadas de Columna revela que no defraudó sólo porque el pago relativo no fuera suficientemente grande. El jugador Columna, además de tratar de maximizar su puntuación, ambicionaba hacer máxima la diferencia entre su puntuación y la de Fila. Pronto manifestó su temor de ser atacado por Fila.

Llegado el momento, el ansia atacante de Columna le hizo llevarse la mejor parte. Su acción dejó desconcertado a Fila, quien pensó que Columna había atacado por miedo a sufrir él el primer ataque o para aumentar el interés del juego. Fila creyó entonces que Columna no volvería a atacarle por miedo a que Fila estuviera planeando vengarse.

Ambos jugadores cooperaron durante los dos turnos siguientes. Fila fue convenciéndose de que Columna no volvería a atacarle, pues él, Fila, había demostrado cuán leal era su conducta. Pero la interpretación que Columna dio a esos acontecimientos era muy distinta. La verdad es que Columna concluyó que Fila estaba solapadamente recomponiendo sus fuerzas, para encontrarse en situación de atacar a Columna con más eficacia. En consecuencia, Columna volvió a atacar en la jugada siguiente. La prolongada cooperación de Fila le estaba intranquilizando: "Quiero llevar sus fuerzas a un nivel tan bajo que si responde a uno de mis ataques quede por debajo de cero".

Después de la partida, Fila reflexionó sobre su estrategia y concluyó que merecía una revisión a fondo. La que ha adoptado subsiguientemente se atiene todavía a principios morales, pero contiene uno de los elementos característicos de la estrategia del ojo-por-ojo: "En esencia consiste en no atacar el primero. En caso de defección del enemigo, tomaré represalias inmediatamente, y así seguiré haciéndolo hasta que mi adversario comience a cooperar o haya dejado absolutamente claro que está dispuesto a ganar a toda costa. En cualquiera de ambos casos yo volveré a cooperar, pues carece de sentido que también yo aniquile a toda su población, aunque él esté decidido a hacerlo con la mía. En cuanto mi oponente deje de cooperar, adoptaré la estrategia de no ser el primero hasta que alcancemos la fase MARV (Fase IV). En esta fase seré yo quien ataque, para eludir el riesgo de sufrir su agresión unilateral, que me dejaría indefenso... Aunque mi análisis pueda parecer cruel, considero que es la mejor forma de asegurar la mínima pérdida de vidas humanas."

A la luz de la experiencia que les proporcionaron las partidas de los estudiantes, Homer-Dixon y Oliveau concluyeron que los jugadores no interpretaban los pagos de la matriz según se les había instruido, sino en función de otros criterios. Los jugadores trataron no pocas veces de hacer máxima la diferencia entre su puntuación y la de su adversario. Ello puede explicar por qué fueron tantos quienes defraudaron nada más empezar. Otros procuraron maximizar la suma de ambos tanteos, lo cual puede explicar por qué un 20 por ciento de los jugadores no defraudó jamás. Hubo todavía jugadores que buscaron la situación de dominancia estratégica, o que, por el contrario, adoptaron posturas puramente éticas.

Uno de los estudiantes resumió su sentir sobre el juego en un ensayo. Citaré algunas de sus conclusiones finales: "Las tentaciones ofrecidas a los jugadores tuvieron fuerza suficiente para inducir a varios estudiantes a lanzarse ataques masivos mutuos... lo cual no se acomoda bien con la teoría de equilibrio de fuerzas, o de equilibrio del terror. Si los dirigentes del mundo comienzan a ver su teatro de operaciones en terminales de vídeo, quizá intervengamos todos en una partida de PACMAN donde acaban devorándose los puntos y la máquina de cambio agota las monedas."

Homer-Dixon cursa el segundo ciclo universitario en el Centro de Estudios Internacionales que el Instituto de Tecnología de Massachusetts tiene en Cambridge (Massachusetts). Oliveau se encuentra en Thinking Machines, Inc., también de Cambridge. La preparación y desarrollo de "After MAD" se hizo con cargo a los presupuestos del Proyecto Atenea, programa que se está desarrollando en el MIT patrocinado por la IBM Corporation y la Digital Research Corporation para estudiar qué papel podrían desempeñar los ordenadores en la enseñanza. Hayward Alker, profesor del departamento de Ciencias Políticas del MIT, contribuyó al desarrollo de "After MAD" y lo ha utilizado en sus cursos. Quien desee obtener más información sobre el juego puede solicitarla a Tad Homer-Dixon, Center for International Studies, 292 Main Street, Cambridge, Mass. 02142, Estados Unidos.

Me ha abrumado la respuesta de los lectores a mi artículo sobre algoblemas del pasado agosto. No era necesario escribir programa alguno,

sino solamente diseñar ciertos algoritmos de repostaje de camiones en el desierto o de cambios de agujas para trenes. Fueron varios centenares las mentes de "vía única" que me enviaron sus más granados logros. Me encantó haber proporcionado algunos vehículos para la reflexión.

Llamé algoblema a aquel tipo de problemas cuya solución es un algoritmo, esto es, una receta o procedimiento inequívoco para lograr un determinado fin. Los problemas de trenes fueron los más sencillos, especialmente habida cuenta de que dí en el artículo la solución de uno de ellos. Como es natural, resultó verdaderamente difícil elegir la mejor solución de entre la multitud de respuestas del problema, que consistía en invertir todas las unidades que componen un tren disponiendo sólo de un ramal de vía con capacidad para un solo vagón. Elegido un poco al azar entre las mejores soluciones que llegaron antes tenemos el siguiente algoritmo, de Douglas A. Owenby, West Covina, California. Como en agosto, la letra *A* denota la sección de vía comprendida entre Problevilla y el ramal; *B* designa el ramal y *C* el tramo de vía que va del ramal a Solucioburgo. El *k*ésimo vagón del tren se denotará *Pk*.

desenganchar el tren *P*
adelante hasta *B*
atrás hasta *C*
para $k = 1$ hasta *n*
  adelante hasta *A*
  enganchar *Pk*
  atrás hasta *C*
  adelante hasta *B*
  desenganchar *Pk*
  atrás hasta *C*
  adelante hasta *A*
  atrás hasta *B*
  enganchar *Pk*

Otra solución parte del último coche del tren, en lugar del primero. Como mencionó buen número de lectores, la solución anterior exige un trabajo total proporcional a n^3, dado que el tren compuesto por *n* vagones habrá de arrastrarse adelante y atrás *n* veces una distancia equivalente a la longitud de *n* vagones.

El primero de los problemas del Zorro del Desierto exigía determinar hasta dónde podría viajar un vehículo de patrulla que dispusiera de *n* bidones de combustible. Inicialmente, el combustible se encuentra en cierto punto del desierto, almacenado en bidones de 200 litros.

El depósito del vehículo tiene capacidad para 40 litros y permite recorrer 200 kilómetros sin repostar; el vehículo puede transportar solamente un bidón cada vez. ¿De qué forma habrá que transportar los bidones por el desierto para hacer máxima la distancia recorrida por el todoterreno? Presenté en el artículo un algoritmo que, para el caso de dos barriles, garantizaba un recorrido de 1200 kilómetros antes de agotar el combustible.

La distancia máxima, como ha demostrado William B. Lipp, de Milford, Connecticut (y otros muchos lectores), es de 1466,66 kilómetros. En una divertida carta, Lipp explica de qué modo pueden superarse los 1200 kilómetros, cosa que yo daba a entender como posible. He aquí su algoritmo (ligeramente retocado):

Llenar el depósito con el primer barril
Cargar el segundo barril
Avanzar 100 kilómetros
Descargar el segundo barril
Regresar a la base
Repostar del primer barril
Cargar el primer barril
Avanzar 200 kilómetros
Descargar el primer barril
Repostar
Regresar 100 kilómetros por el segundo barril
Cargar el segundo barril
Avanzar 100 kilómetros hasta el primer barril
Repostar del primer barril
Avanzar $66\frac{2}{3}$ kilómetros
Descargar el segundo barril
Regresar $66\frac{2}{3}$ kilómetros hasta el primer barril
Cargar el primer barril
Avanzar $66\frac{2}{3}$ kilómetros hasta el segundo barril

En ese momento, el vehículo se encuentra a $266\frac{2}{3}$ kilómetros del punto de partida. Reposta por última vez del primer barril. Carga el segundo barril (que sigue lleno) y se interna otros 1200 kilómetros en el desierto.

Este algoritmo no es susceptible de generalización a un algoritmo óptimo cuando hay más de $n = 2$ barriles. Por ejemplo, no permite alcanzar los 1720 kilómetros correspondientes a la solución óptima en el caso de $n = 3$ barriles.

Daré el mes que viene un algoritmo general (para el caso de $n \geq 6$ barriles) y presentaré asimismo una solución para el segundo de los problemas del Zorro del Desierto.

Libros

Etica médica, osciladores biológicos, ciencia del suelo e invertebrados

Carlos M. Romeo, David Jou, Juan Gallardo y Rafael Jordana

ETICA MÉDICA. LAS RESPONSABILIDADES MORALES DE LOS MÉDICOS, por Tom L. Beauchamp y Laurence B. McCullough. Labor; Barcelona, 1987. Los autores de esta obra pertenecen a la Universidad de Georgetwon y al Instituto Kennedy de Etica de la misma, uno de los centros más prestigiosos y cualificados de investigación sobre temas relacionados con la ética general y con la bioética.

El título es claramente revelador de su contenido: se ocupa de las responsabilidades morales de los médicos en su asistencia clínica y, en concreto, aporta unas pautas decisionales para ayudar a resolver los conflictos de índole moral a los que han de enfrentarse aquéllos en su quehacer diario, desde los más rutinarios hasta los más complejos. Dividida en seis capítulos, en el primero de ellos se nos muestra cómo opera el razonamiento moral en medicina; el segundo describe los principios o "modelos" básicos (principios de beneficencia y de autonomía) para ayudar a la toma de decisiones morales en el ejercicio de la medicina; los cuatro restantes contrastan la validez de aquellos principios en situaciones especialmente delicadas, aunque no por ello inhabituales, que sirven de piedra de toque para explicar su funcionamiento: el manejo de la información médica, el llamado paternalismo médico, los supuestos de autonomía reducida o capacidad disminuida del paciente y los intereses de terceras partes (familiares y empresarios del paciente, poderes públicos, etcétera). Todas ellas parten de la indagación de los conflictos que pueden surgir y exponer los mecanismos de su resolución a la luz de los principios aplicables.

Nos encontramos, sin duda, ante una obra de particular interés, no sólo por la actualidad del tema abordado, en el que los autores poseen reconocida experiencia, sino, sobre todo, por el peculiar enfoque con que es tratado y resuelto. En efecto, el lector se verá sorprendido por la metodología utilizada, tan querida para los pensadores anglosajones, como es el método del caso. Esto significa que la discusión gira en torno a los casos propuestos, al hilo de los cuales van construyendo los autores los rasgos más característicos de los principios que les sirven de referencia, así como su sustento filosófico e histórico y su fundamento moral; de esta manera se está en condiciones de intentar resolver a continuación el conflicto o conflictos morales subyacentes y de tomar la decisión lo más correcta y racional posible. Este procedimiento presenta enormes ventajas. Por un lado, se va introduciendo sin esfuerzo al lector en el núcleo de la cuestión, al tener siempre la referencia de la situación real planteada, que él mismo ha podido vivir (si es un profesional de la medicina) o conocer en su experiencia personal; por otro, va pasando del pensamiento concreto al abstracto sin apenas notarlo, lo que aligera la lectura; pero, sobre todo, permite al lector situarse no como un mero receptor de información, que podrá aceptar o no, sino como parte activa de la obra junto a sus autores, pues va tomando postura, casi sin querer, frente a cada una de las opciones de solución que se le van ofreciendo, al ser fácilmente contrastables con el caso que sirve de punto de partida. Este último mecanismo se ve favorecido por la misma libertad que propician los autores, al no asumir compromisos radicales, sino razonados y abiertos a la discrepancia y la crítica, sin que ello quiera significar que sus posturas no sean en cada momento claras. Late ahí la idea del respeto a las convicciones de los demás en un clima de tolerancia, actitud irrenunciable en una sociedad pluralista.

De ahí el gran valor pedagógico y formativo de la obra, que ha de ser de gran utilidad para estudiantes de medicina, médicos e interesados en temas morales, en cuanto desborda con sus planteamientos el campo de la medicina, aunque se centre de modo constante en ella. Pues, como apunta certeramente el prologuista, Diego Gracia, hay que abandonar la vieja idea de que estos asuntos no necesitan aprenderse, de que basta el sentido común y un poco de prudencia; muy al contrario, la toma de decisiones morales en el ámbito sanitario es enormemente compleja y requiere un entrenamiento específico.

¿Qué hacer ante un neonato inien-cefálico? ¿o cuando una mujer adulta, pero joven y capaz solicita ser sometida a una esterilización quirúrgica? ¿hasta dónde debe llegar el médico en la información al paciente, sobre todo si se trata de una persona psicológicamente vulnerable o padece de una enfermedad terminal? ¿y si en esta situación el paciente solicita con plena conciencia que se le desconecte el respirador mecánico? A estas situaciones y otras muchas más hay que dar una respuesta adecuada. Pero para un razonamiento adecuado son imprescindibles dos premisas: la claridad, de forma que podamos identificar los conceptos básicos referentes a la situación, eliminando en lo posible las equivocaciones y ambigüedades, y la coherencia, con el fin de que nuestro razonamiento esté libre de contradicciones.

Y es así como pueden entrar en juego esos principios o modelos de *autonomía* y de *beneficencia,* que dan lugar a determinados deberes en el médico. De acuerdo con el primero, hay que aceptar que las demás personas se autogobiernan correctamente (son autónomas), tanto en sus elecciones como en sus actos, mientras que conforme al segundo debemos proporcionar cosas beneficiosos y prevenir o suprimir las perjudiciales. Estos dos principios coinciden en que ambos persiguen el objetivo moral de procurar "el mejor interés del paciente", sólo que el de autonomía lo concibe desde la perspectiva del paciente, de cómo lo entienda él mismo, lo que implica el respeto del médico hacia los valores y creencias de aquél; mientras que el de beneficencia orienta ese mejor interés del paciente desde el punto de vista de la medicina, con lo que se logra una perspectiva objetiva, al trascender los valores y creen-

cias de cada médico concreto. De todas formas, cada uno de estos modelos ofrece una perspectiva válida, pero parcial, de las responsabilidades de los médicos; sin embargo, cada uno de ellos fundamenta lo que se denomina un deber *prima facie*, como deber que siempre hay que cumplir, salvo que entre en conflicto en algún caso con un deber igual o más fuerte. A orientar sobre estas situaciones ayudan también los autores. Para determinados supuestos puede entrar también en juego el —principio del— *paternalismo*, basado en la preparación técnica y mayor experiencia del médico, particularmente cuando la capacidad de decisión autónoma del paciente está limitada por cualquier causa permanente o meramente circunstancial. Sin embargo, su aplicación no deja de ser discutida y aún discutible.

Una vez realizado (o mejor: a la vez que) el análisis moral, hay que contrastarlo con las prescripciones del derecho. Si existe coincidencia, el camino a la actuación estará abierto. En caso contrario (y esto podríamos añadirlo a la vista de nuestro derecho), podrá alegarse la objeción de conciencia, salvo que la no intervención ponga en peligro la vida o la salud del paciente.

En conclusión, si a este prometedor panorama añadimos la mención de que se incluye una seleccionada bibliografía de apoyo o de contraste, podemos afirmar que nos encontramos ante una obra de aconsejable lectura, si no necesaria, para cuantos se mueven en el mundo de la medicina y han de tomar decisiones morales responsables y para quienes se están preparando para ello. (C.M.R.)

TEMPORAL DISORDER IN HUMAN OSCILLATORY SYSTEMS, dirigido por L. Rensing, U. van der Heiden y M.C. Mackey. Springer Series in Synergetics; Springer Verlag; Berlín, 1987. Suficientemente lejos del equilibrio, los sistemas macroscópicos descritos por leyes no lineales tienden a ordenarse. Esta ordenación conduce a una autoestructuración espontánea, de carácter espacial o temporal. Resulta sorprendente que en sistemas muy diversos —hidrodinámicos, ópticos, químicos, biológicos, geológicos, etcétera— el proceso de esta autoorganización presenta ciertas características comunes, más o menos análogas a las de las transiciones de fase en equilibrio. Fue Prigogine el primero en destacar el interés científico de esta regularidad, en acuñar el nombre y el concepto de *estructuras disipativas* para tales organizaciones y en atraer científicos hacia este campo. Posteriormente, Haken, hacia 1975, dio nuevo impulso a este interés, lo sistematizó y lo denominó *sinergética*, o ciencia de la autoorganización colectiva.

Desde entonces, a través de una serie de congresos y de publicaciones, la sinergética ha conocido un auge creciente. Así, el volumen que comentamos es ya el trigésimo sexto de la colección que Springer le ha dedicado. Aunque la gran mayoría de los estudios agrupados en dicha colección se hubieran llevado igualmente a cabo sin la creación de la sinergética, este intento organizativo ha resultado ciertamente fructífero en su difusión y popularización y ha fomentado activamente la interdisciplinariedad.

En el presente volumen se agrupa una colección de veintiocho ponencias presentadas en un simposio internacional llevado a cabo en la Universidad de Bremen en septiembre de 1986. Su temática gira, como ya indica el título, alrededor del estudio de los diversos ritmos del cuerpo humano (en los sistemas respiratorio, circulatorio, nervioso, neuromotor y reproductor, principalmente) y, en especial, de sus interacciones y de sus alteraciones.

En algunas ocasiones, éstas se producen como resultado normal del desarrollo del individuo (desarrollo de ritmos circadianos en recién nacidos, aparición de ciclos hormonales en la pubertad y su desaparición tras la menopausia). En otros casos, suponen enfermedades de diverso grado de gravedad: taquicardias y fibrilaciones ventriculares en el corazón, temblores incontrolados en el sistema neuromotor, algunos tipos de epilepsia en trastornos cerebrales, diversas clases de depresiones al variar el ritmo circadiano (de periodicidad aproximada de 24 horas) del sueño y la vigilia. Finalmente, algunas enfermedades presentan, asimismo, ciertos ciclos característicos, cuyo estudio es indispensable para su tratamiento. Así, se ha llegado a hablar incluso de *enfermedades dinámicas*, que son las que se producirían en un sistema fisiológicamente intacto debido a trastornos puramente dinámicos. El objetivo final de los autores es la comprensión de los diversos ritmos del cuerpo (su origen, su interacción, sus diferencias de fase) para poder hallar la terapéutica más adecuada en cada caso.

El libro está dividido en cinco partes, una dedicada a conceptos teóricos básicos y las otras cuatro referidas a distintos sistemas y órganos: nervioso y neuromotor, corazón y respiración, ritmos circadianos y ciclos ováricos. Algunas cuestiones son resultado directo de progresos recientes en la comprensión de sistemas dinámicos no lineales, y el uso de algunos conceptos de este dominio (dimensiones fractales, atractores extraños, exponentes de Lyapunov) aparece en diversos artículos. Ejemplo de ello es la interpretación de electroencefalogramas (Babloyantz y Destexhe, Mayer-Kreiss y Holzfuss). ¿Deben interpretarse éstos como una suma de muchos osciladores lineales débilmente acoplados, o como unos pocos osciladores no lineales fuertemente acoplados? ¿Cuál es el tamaño de las regiones cerebrales de actividad electromagnética coherente? ¿Qué características del encefalograma pueden dar alguna pista sobre estas cuestiones?

El uso de técnicas de dinámica no lineal se encuentra también en el estudio de diversas alteraciones cardíacas (taquicardia, fibrilación ventricular, infarto de miocardio). En este campo se plantean aún disyuntivas muy serias: ¿es la fibrilación ventricular un caso típico de caos dinámico o turbulencia como se creía hasta el momento? ¿Es, por el contrario, un caso de periodicidad patológica? A esto último apuntan, en contra de la opinión al uso, algunos experimentos recientes (Goldberger).

En otros casos, el tratamiento del problema es más cualitativo, muy poco influido por las técnicas mencionadas. Esta es la tendencia general del apartado dedicado a ritmos circadianos, centrado en las consecuencias psiquiátricas (depresiones, insomnios graves) que conllevan las alteraciones en la periodicidad sueño-vigilia, debidas bien a variaciones estacionales en la duración de la fase luminosa del día, bien a la realización de trabajos nocturnos. Se examinan diversos tipos de tratamiento, en especial los basados en el efecto de la luz brillante sobre la segregación de melatonina, sustancia levemente sedante segregada por el hipotálamo, en especial durante la noche, y que puede ser parcialmente inhibida cuando la luminosidad ambiental sobrepasa ciertos límites. Algunas preguntas técnicas abiertas en este campo se refieren al número de osciladores necesarios para describir la periodicidad en la segregación de melatonina: ¿basta un oscilador o se requieren dos?

La última sección, sobre ciclos ováricos, es de carácter más matemático y estudia la relación entre el número de folículos excitados en cada ciclo ovárico con la regularidad numérica (óvulos madurados en cada ciclo) y temporal

de la ovulación. Según los resultados del modelo, la activación de un gran número de folículos, la mayoría de los cuales no llegan a madurar, favorece la regularidad en estas características.

En resumen, el libro es actual e interesante, pero es poco sistemático y resulta poco adecuado como introducción al tema. La voluntad de los autores ha preferido exponer avances recientes para el especialista en el campo concreto que optar por proporcionar una visión general y pedagógica del tema. (D.J.)

MANUAL DE EDAFOLOGÍA, por Ph. Duchaufour. Masson; Barcelona, 1987. Acaba de aparecer este texto con tres años de retraso respecto al original francés. Recuerda el clásico libro del mismo autor que servía de base para obtener el diploma de estudios profundos en pedología ("Processus de formation des sols: Biochimie et Géochimie"; Nancy, 1972), sobretodo en la exposición de temas, aun cuando se observan algunas diferencias derivadas, sin duda, de la acostumbrada continua corrección que el prolífico profesor Duchaufour suele hacer de sus obras.

Así, el manual se inicia abordando los constituyentes edáficos, separando la fracción mineral y la fracción orgánica; de la primera destaca el complejo de alteración (arcillas y oxihidróxidos) y de la segunda el proceso de humificación. Sin embargo, se echa en falta el capítulo diferenciado de la formación y evolución de los complejos organominerales que daba una visión integrada de los componentes del suelo en el antiguo libro citado.

La segunda parte versa sobre las propiedades del suelo, fundamentalmente las físicas (estructura, agua edáfica, temperatura) y fisicoquímicas (complejo absorbente), para luego introducir al lector en el concepto de fertilidad del suelo.

Se pasa así a la tercera parte, que comienza repasando los factores formadores y procesos de evolución del suelo, destacándose de manera relevante los procesos derivados de la incorporación de materia orgánica al suelo originándose, en consecuencia, horizontes A (carbonatación, empardecimiento, lixiviación, podsolización), así como los derivados de la alteración del roquedo, con la consiguiente formación de arcillas y de los horizontes B (fersialitización, ferruginación, ferralitización); por tanto, quedan incluidos los procesos de tercer tipo (los que implican redistribución de materia) en uno u otro de los grupos anteriormente citados. A continuación se entra de pleno en la sistemática de suelos que, siendo genética (basada en procesos biogeoquímicos), abandona criterios ecológicos que son utilizados por la clasificación francesa "oficial". De ese modo, se centra en caracteres propios de los procesos edafogenéticos; esto conduce a la definición de clases (y subclases): I. Suelos pocos evolucionados climáticos (criosoles y suelos subdesérticos), de erosión (litosoles y regosoles) o de aporte (suelos aluviales y coluviales); II. Suelos poco diferenciados (ránkeres y andosoles); III. Suelos calcimagnésicos (rendsinas, suelos calcimagnésicos empardecidos y humificados); IV. Suelos empardecidos, de alteración (suelos pardos y ocres), lixiviados (suelos pardos lixiviados) y lixiviados continentales (suelos dernovopodsólicos y grises forestales); V. Suelos podsolizados (suelos podsólicos, podsoles y suelos podsolizados hidromorfos); VI. Vertisoles; VII. Suelos isohúmicos (chernozem, castanozem, brunizem, suelos marrones y semiáridos); VIII. Suelos fersialíticos; IX. Suelos ferruginosos; X. Suelos ferralíticos; XI. Suelos hidromorfos fuertemente gleizados (pseudogleis, estañogleis, gleis y turbas) o con gleización reducida (pelosoles y planosoles), y XII. Suelos salsódicos (salinos y alcalinos).

La cuarta y última parte contiene nociones de cartografía y posibles aplicaciones de los conocimientos expuestos, que da al lector una sensación final de utilidad de la ciencia del suelo y, por ello, acrecienta su posible interés.

Una objeción a realizar es la fácil tendencia a traducir "pédologie" por "edafología", no siendo exactamente equivalente; al igual que en la introducción el mismo Duchaufour distingue "ciencia del suelo" (aspecto estático) de la "pedología" (aspecto dinámico), la disciplina "edafología" se diferencia de esta última porque la edafología versa, además, sobre la relación del suelo con la planta; es cierto que el término pedología no es tan popular en España como en Francia, pero ello no justifica la extensión del equívoco. Por otra parte, se mantiene del original la unidad "pulgadas" en alguna figura, lo cual podía haber sido fácilmente convertido en centímetros y, así, facilitar el significado al lector habituado al sistema métrico decimal.

De todas formas, creemos que el suscrito *Manual de edafología* (o, mejor, de pedología) rellena un hueco editorial y se encuentra a caballo entre la alta divulgación y el libro-guía del estudiante de primer ciclo universitario;

en todo caso, se dirige a un público con una sólida formación básica. Finalmente, hay que elogiar el esfuerzo que vienen realizando los doctores Carballas en verter al castellano la importante y abundante obra escrita del profesor Duchaufour. (J.G.)

ZOOLOGÍA DE INVERTEBRADOS, por P.A. Meglitsch. Editorial Pirámide; Madrid, 1986. La zoología es una ciencia muy antigua, y hay muchos y muy buenos libros para su estudio. La obra que nos ocupa es la segunda edición en España —la primera la hizo la editorial Blume—, traducida de la segunda edición inglesa de 1972. No es, por tanto, una novedad en el mercado, pero el hecho de su reimpresión es un índice de la aceptación de esta zoología general de invertebrados.

La obra se estructura en 19 capítulos y tres apéndices. Sigue un orden sistemático por niveles de organización desde los protozoos a los deuteróstomos. Los diferentes *phyla* son tratados agrupados en cada capítulo con aquellos otros con los que muestran relaciones filogenéticas o similaridades embrionarias.

Cuando un autor aborda la tarea de escribir un libro de carácter general, necesariamente vierte el enfoque personal que tiene de esa área de la ciencia, si bien debe estructurar, adecuar, recortar, etcétera, con un análisis minucioso, los conocimientos adquiridos sobre esa disciplina, para adaptarlos al fin propuesto. En este caso, Meglitsch ha realizado una verdadera introducción a la zoología general de los invertebrados muy equilibrada y completa. De las 900 páginas que ocupa el texto, 450 las dedica a los invertebrados no artrópodos; 233 al *phylum* de los artrópodos y el resto a los celomados deuterostomos y grupos afines. Esta distribución parece suficientemente equilibrada, porque se han tenido en cuenta el gran número de tipos diferentes de organización entre los invertebrados no artrópodos y a la vez se les da extensión suficiente a las distintas clases de artrópodos, extensión necesaria por la gran radiación adaptativa que se observa en este *phylum* y por el elevado número de especies que lo forman.

Por otro lado, la estructuración de cada capítulo es la idónea. En todos y cada uno de los *phyla* se sigue el esquema siguiente: *a*) Se explican de un modo ordenado las características más notables, lo que facilita su comprensión. *b*) Tras su clasificación, se describe cada uno de los grupos taxonómicos. El grado de profundidad en la clasificación propuesta depende de la complejidad del grupo, pero en general se llega a la descripción de los diferentes órdenes de cada clase. *c*) Se realiza una explicación funcional de la morfología. No es, en efecto, la estructura lo que interesa en la anatomía de los animales, sino la función que realiza esa estructura en la unidad funcional del individuo. Es ésta la aportación más importante: una visión funcional de la anatomía, o mejor dicho, una explicación morfo-funcional.

Es indudable que, de un libro de estas características, no cabe esperar un tratamiento complejo de la fisiología animal comparada, y mucho menos mantenerla al día. La fisiología es una ciencia experimental, y en cuanto tal, abierta a un aumento constante de los conocimientos, mientras que la zoología avanza a un ritmo más pausado. Se da también otra circunstancia: las interpretaciones funcionales cambian ante nuevos descubrimientos, y los textos y monografías sobre fisiología pierden su vigencia muy pronto. Esto no sucede con la zoología, que, al tratar de los aspectos morfológicos, anatómicos y sistemáticos, disfruta de mayor permanencia.

En la "biología" de cada grupo suelen estudiarse los efectores, aparato digestivo, tipos de alimentación, respiración, circulación, órganos de los sentidos, sistema nervioso, reproducción y desarrollo embrionario. Por tanto, la visión que se da de la zoología no es simplemente anatómica, estática, sino que recibe un enfoque anatomo-funcional, dinámico e integrado en el medio donde vive.

La sistemática zoológica utilizada resulta algo anticuada, resintiéndose de la fecha de la edición (1972) que ha servido para la traducción, si bien, por ser clásica y descender sólo hasta el nivel de orden, es perfectamente válida y didáctica. Son de destacar la gran profusión de ilustraciones, dibujos semiesquemáticos y fotografías que facilitan la comprensión del texto. Estas ilustraciones tienen extensos pies de figura, independientes del texto; los nombres técnicos de los diferentes detalles morfológicos de los dibujos están en el propio dibujo.

Los dos primeros capítulos (introducción y criterios sobre nomenclatura zoológica) revisten especial interés para los no iniciados; sirven de gran ayuda los apéndices finales sobre filogenia animal, estratificación de hábitats marinos y el resumen de la historia geológica. Se echa en falta un glosario de términos científicos (R.J.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

LA PROXIMA REVOLUCION INFORMATICA

NO SILVER BULLET: ESSENCE AND ACCIDENTS OF SOFTWARE ENGINNERING. Frederick P. Brooks, Jr. en *Computer*, vol. 20, n.º 4, págs. 10-19; abril, 1987.

HERE COME THE BILLION-TRANSISTOR IC. Bernard C. Cole en *Electronics*, vol. 60, n.º 7, págs. 81-85; 2 de abril de 1987.

ARQUITECTURAS AVANZADAS DE COMPUTADORES

PARALLEL COMPUTERS. R. W. Hockney y C. R. Jesshope. Adam Hilger, Ltd., 1981.

A SURVEY OF HIGHLY PARALLEL COMPUTING. Leonard S. Haynes, Richard L. Lau, Daniel P. Slewiorek y David W. Mizell en *Computer*, vol. 15, n.º 1, págs. 9-24, enero, 1982.

CONCURRENT VLSI ARCHITECTURES. Charles L. Seitz en *IEEE Transactions on Computers*, vol. C-33, n.º 12, págs. 1247-1265; diciembre, 1984.

A SURVEY OF PROPOSED ARCHITECTURES FOR THE EXECUTION OF FUNCTIONAL LANGUAGES. Steven R. Vegdahl en *IEEE Transactions on Computers,* vol. C-33, n.º 12, págs. 1050-1071; diciembre, 1984.

EXPERIMENTAL PARALLEL COMPUTING ARCHITECTURES. Dirigido por Jack C. Dongarra. Elsevier Science Publishing Co., Inc., 1987.

PROCEEDING OF THE THIRD CONFERENCE ON MULTIPROCESSORS AND ARRAY PROCESSORS. Dirigido por Walter J. Karplus. The Society for Computer Simulation, 1987.

MICROCIRCUITOS PARA COMPUTACION AVANZADA

THE EVOLUTION OF DIGITAL ELECTRONICS TOWARDS VLSI. Robert W. Keyes en *IEEE Transactions on Electron Devices*, vol. ED-26, n.º 4, págs. 271-279; abril, 1979.

ULTRA-LARGE SCALE INTEGRATION. James D. Meindl en *IEEE Transactions on Electron Devices,* volumen ED-31, n.º 11, págs. 1555-1561; noviembre, 1984.

PROGRAMAS PARA COMPUTACION AVANZADA

THE COSMIC CUBE, Charles L. Seitz en *Communications of the ACM*, volumen 28, número 1, páginas. 22-23; enero, 1985.

DOMESTICATING PARALLELISM. David Gelernter en *Computer*, vol. 19, n.º 8, págs. 12-19; agosto, 1986.

PARA-FUNCTIONAL PROGRAMMING. Paul Hudak en *Computer*, vol. 19, n.º 8, págs. 60-71; agosto, 1986.

ADVANCED COMPUTER OPTIMIZATIONS FOR SUPERCOMPUTERS. David A. Padua y Michael J. Wolfe en *Communications of the ACM*, vol. 29, n.º 12, págs. 1184-1201; diciembre, 1986.

DATA PARALLEL ALGORITHMS. W. Daniel Hillis y Guy L. Steele, Jr., en *Communications of the ACM*, vol. 29, n.º 12, págs. 1170-1183; diciembre, 1986.

PARALLELISM, PERSISTENCE AND METACLEANLINESS IN THE SYMMETRIC LISP INTERPRETER. David Gelernter, Suresh Jaganathan y Thomas London en *Proceedings of the SIGPLAN'87 Symposium on Interpreters and Interpretive Techniques*. ACM Press, 1987.

ALMACENAMIENTO DE DATOS PARA COMPUTACION AVANZADA

TECNOLOGIAS DE LAS MEMORIAS DE DISCO. Robert M. White en *Investigación y Ciencia*, n.º 49, págs. 94-104; octubre, 1980.

PERPENDICULAR MAGNETIC RECORDING-EVOLUTION AND FUTURE. Shunichi Iwasaki en *IEEE Transactions on Magnetics*, vol. MAG-20, n.º 5, págs. 657-662; septiembre, 1984.

MAGNETO-OPTIC RECORDING TECHNOLOGY. Mark H. Kryder en *Journal of Applied Physics*, vol. 57, n.º 8, parte IIB, págs. 3913-3918; 15 de abril de 1985.

INTERFASES PARA ORDENADORES AVANZADOS

FUNDAMENTALS OF INTERACTIVE COMPUTER GRAPHICS. James D. Foley y Andries van Dam. Addison-Wesley Publishing Co., 1982.

DIRECT MANIPULATION: A STEP BEYOND PROGRAMMING LANGUAGES. Ben Shneidermann en *Computer*, vol. 16, n.º 8, págs. 57-69; agosto, 1983.

THE HUMAN INTERFACE: WHERE PEOPLE AND COMPUTERS MEET. Richard A. Bolt. Lifetime Learning Publications, 1984.

DIRECT MANIPULATION INTERFACES. E. Hutchins, J. Hollan y D. Norman en *User Centered System Design*, dirigido por Donald A. Norman y Stephen Draper. Lawrence Erlbaum & Associates, 1986.

AVANCES EN REDES DE ORDENADORES

COMPUTER NETWORK DEVELOPMENT TO ACHIEVE RESOURCE SHARING. L. G. Roberts y B. D. Wessler en *Proceedings of the AFIPS Spring Joint Computer Conference.* American Federation of Information Processing Societies, 1970.

A PROTOCOL FOR PACKET NETWORK INTERCOMMUNICATION. Vinton G. Cerf y Robert E. Kahn en *IEEE Transactions on Communications*, vol. COM-22, n.º 5, págs. 637-648; mayo, 1974.

SPECIAL ISSUE ON PACKET COMMUNICATION NETWORKS. *Proceedings of the IEEE*, vol. 66, n.º 11; noviembre, 1978.

EVOLUTION OF THE ETHERNET LOCAL COMPUTER NETWORK. John F. Shoch, Yogen K. Dalal, David D. Redell y Ronald C. Crane en *Computer*, vol. 15, n.º 8, págs. 10-27; agosto, 1982.

INTEGRATED SERVICES DIGITAL NETWORK. *AT&T Technical Journal*, vol. 65, n.º 1; enero-febrero, 1986.

COMPUTACION AVANZADA EN EL CAMPO DE LA CIENCIA

GALACTIC BRIDGES AND TAILS. Alar Toomre y Juri Toomre en *The Astrophysical Journal,* vol. 178, n.º 3, parte 1, págs. 623-666; 15 de diciembre de 1972.

INTRODUCTION TO VLSI SYSTEMS. Carver Mead y Lynn Conway. Addison-Wesley Publishing Co., 1980.

THE PHYSICAL UNIVERSE: AN INTRODUCTION TO ASTRONOMY. FRANK H. Shu. University Science Books, 1982.

BINARY FORMATION AND INTERACTIONS WITH FIELD STARS. Piet Hut en *Dynamics of Star Clusters: Proceedings of the 113th Symposium of the International Astronomical Union*, dirigido por Jeremy Goodman y Piet Hut. D. Reidel Publishing Co., 1985.

A DIGITAL ORRERY, James H. Applegate, Michael R. Douglas, Yetka Gürsel, Gerald J. Sussman y Jack Wisdom en *The Astronomical Journal*, vol. 92, n.º 1, págs. 176-194; julio, 1986.

COMPUTACION AVANZADA EN EL CAMPO DE LA MEDICINA

READINGS IN MEDICAL ARTIFICIAL INTELLIGENCE: THE FIRST DECADE. Dirigido por W. J. Clancey y E. H. Shortliffe. Addison-Wesley. Publishing Co., 1984.

COMPUTER-ASSISTED MEDICAL DECISION MAKING. Dirigido por J. A. Reggia y S. Tuhrim. Springer-Verlag, 1985.

A COMPUTATIONAL MODEL OF REASONING FROM THE CLINICAL LITERATURE. Glenn D. Rennels. Springer-Verlag, 1985.

SPECIAL ISSUE ON MEDICAL INFORMATICS. *The Western Journal of Medicine,* vol. 145, n.º 6; junio, 1986.

A COMPUTER DATA BASE FOR INFORMATION ON CANCER TREATMENT. Susan Molloy Hubbard, Jane E. Henney y Vincent T. DeVita en *The New England Journal of Medicine*, vol. 316, n.º 6, págs. 315-318; 5 de febrero de 1987.

COMPUTER PROGRAMS TO SUPPORT CLINICAL DECISION MAKING. Edward H. Shortliffe en *The Journal of the American Medical Association*, vol. 258, n.º 1, págs. 61-66, 3 de julio de 1987.

COMPUTACION AVANZADA EN LA INDUSTRIA

IMPACT OF COMPUTATIONAL METHODS ON AIRCRAFT DESIGN. P. E. Rubbert y E. N. Tinoco. AIAA Atmosphere Flight Conference, 1983.

HIGH-SPEED COMPUTATION. Dirigido por Janusz S. Kowalik. Springer-Verlag. 1984.

VECTOR AND PARALLEL PROCESSORS IN COMPUTATIONAL SCIENCE. Dirigido por I. S. Duff y J. K. Reid. Elsevier Science Publishing Co., Inc., 1985.

DIRECT METHODS FOR SPACE MATRICES. I. S. Duff, A. M. Erisman y J. K. Reid. Oxford University Press, 1986.

JUEGOS DE ORDENADOR

FIGHTS, GAMES, AND DEBATES. Anatol Rapoport. University of Michigan Press, 1960.

THE EVOLUTION OF COOPERATION. Robert Axelrod. Basic Books, 1984. (Traducción española de Alianza Universidad, Madrid, 1986.)

Indice anual

Este índice contiene los nombres de los autores y artículos aparecidos en Investigación y Ciencia *a partir de enero de 1987. Se incorpora asimismo la relación alfabética de las secciones* Juegos de ordenador *y* Taller y laboratorio.

AUTORES

Abbott, Allan V., Alec N. Brooks y D. Gordon Wilson. HIDRONAVES DE TRACCIÓN HUMANA: febrero, pág. 76.

Abu-Mostafa, Yaser S., y Demtri Psaltis. COMPUTADORAS OPTICONEURONALES; mayo, pág. 58.

Ada, Gordon L., y Sir Gustav Nossal. LA TEORÍA DE LA SELECCIÓN CLONAL; octubre, pág. 40.

Allen, Robert Day. EL MICROTÚBULO, MOTOR INTRACELULAR; abril, pág. 8.

Appenzeller, Tim, y Mortimer Mishkin. ANATOMÍA DE LA MEMORIA; agosto, pág. 15.

Barón, Matilde, Ana Chueca y J. López Gorgé. INHIBICIÓN DE LA FOTOSÍNTESIS POR HERBICIDAS; julio, pág. 10.

Barrett, Spencer C.H. MIMETISMO VEGETAL; noviembre, pág. 58.

Baylor, Denis A., y Julie L. Schnapf. RESPUESTA DE LOS FOTORRECEPTORES A LA LUZ; junio, pág. 20.

Beck, Gregory, Gail S. Habicht y Jorge L. Benach. LA ENFERMEDAD DE LYME; septiembre, pág. 52.

Beek, Gus W. Van. ARCOS Y BÓVEDAS DEL PRÓXIMO ORIENTE; septiembre, pág. 76.

Benach, Jorge L., Gail S. Habicht y Gregory Beck. LA ENFERMDAD DE LYME; septiembre, pág. 52.

Bernáldez, F.G., P. Herrera, C. Levassor, B. Peco y A. Sastre. LAS AGUAS SUBTERRÁNEAS EN EL PAISAJE; abril, pág. 8.

Bharathan, Desikan, y Terry R. Penney. ENERGIA EXTRAÍDA DEL MAR; marzo, pág. 58.

Blakeway, John, y Martine Seu-Salerno. EL PERFUME; septiembre, pág. 38.

Bloembergen, Nicolaas, y C. Kumar N. Patel. DEFENSA ESTRATÉGICA Y ARMAS DE ENERGÍA DIRIGIDA; noviembre, pág. 14.

Bonatti, Enrico. LA ROTURA DE LOS CONTINENTES; mayo, pág. 66.

Brooks, Alec N., Allan V. Abbott y D. Gordon Wilson. HIDRONAVES DE TRACCIÓN HUMANA; febrero, pág. 76.

Brown, Robert Hamilton, Torrence V. Johnson y L. A. Soderblom. LOS SATÉLITES DE URANO; junio, pág. 30.

Brownlee, Michael, Anthony Cerami y Hellen Vlassara. GLUCOSA Y ENVEJECIMIENTO; julio, pág. 52.

Campillo, Mercedes, Serafín Fraga, Luis Seijo y Bhagirath Singh. PREVENCIÓN DEL RECHAZO DE TRANSPLANTES; enero, pág. 6.

Campos-Ortega, José A. y Fernando Jiménez. BASES DE LA NEUROGÉNESIS EN *DROSOPHILA MELANOGASTER*; septiembre, pág. 8.

Cárdenas, Jacobo y Francisco Castillo. FIJACIÓN BIOLÓGICA DEL NITRÓGENO; noviembre, pág. 88.

Carter. William E., y Douglas S. Robertson. GEOLOGÍA POR INTERFEROMETRÍA DE MUY LARGA BASE; enero, pág. 30.

Castillo, Francisco, y Jacobo Cárdenas. FIJACIÓN BIOLÓGICA DEL NITRÓGENO; noviembre, pág. 88.

Cech, Thomas R. FUNCIÓN ENZIMÁTICA DEL ARN; enero, pág. 42.

Cerami, Anthony, Helen Vlassara y Michael Brownlee. GLUCOSA Y ENVEJECIMIENTO; julio, pág. 52.

Cole, Jonathan R., y H. Zuckerman. RENDIMIENTO DE LA MUJER CASADA EN LA INVESTIGACIÓN; abril, pág. 76.

Crutchfield, James P., J. Doyne Farmer, Norman H. Packard y Robert S. Shaw. CAOS; febrero, pág. 16.

Cuzzi, Jeffrey N., y Larry W. Esposito. LOS ANILLOS DE URANO; septiembre, pág. 18.

Chávez, Francisco. EL NIÑO Y LA OSCILACIÓN DEL SUR; mayo, pág. 46.

Childress, James J., Horts Felbeck y G. N. Somero. SIMBIOSIS EN LAS PROFUNDIDADES MARINAS; julio, pág. 78.

Chow, Marie, James M. Hogle y David J. Filman. ESTRUCTURA DE LOS POLIOVIRUS; mayo, pág. 22.

Chueca, Ana, Matilde Barón y J. López Gorgé. INHIBICIÓN DE LA FOTOSÍNTESIS POR HERBICIDAS; julio, pág. 10.

Davis, William C. DETONACIÓN DE EXPLOSIVOS; julio, pág. 70.

DeVries, Arthur L., y J. T. Eastman. PECES ANTÁRTICOS; enero, pág. 60.

Dressler, A. EL GRANDIOSO DESFILE DE LAS GALAXIAS; noviembre, pág. 22.

Dye, James L. ELECTRUROS; noviembre, pág. 32.

Eastman, Joseph T., y Arthur L. DeVries. PECES ANTÁRTICOS; enero, página 60.

Eastman, Lester F., y Mordehai Heiblum. ELECTRONES BALÍSTICOS EN SEMICONDUCTORES; abril, pág. 56.

Erisman, Albert M., y Kenneth W. Neves. COMPUTACIÓN AVANZADA EN LA INDUSTRIA; diciembre, pág. 118.

Esposito, Larry W., y Jeffrey N. Cuzzi. LOS ANILLOS DE URANO; septiembre, pág. 28.

Evans, Michael L., Randy Moore y Karl-Keinz Hasenstein. RESPUESTA DE LAS RAÍCES A LA GRAVEDAD; febrero, pág. 68.

Farmer, J. Doyne, James P. Crutchfield, Norman H. Packard y Robert S. Shaw. CAOS; pág. 16.

Felbeck, Horts, James J. Childress y G. N. Somero. SIMBIOSIS EN LAS PROFUNDIDADES MARINAS; julio, pág. 78.

Filman, David J., James M. Hogle y Marie Chow. ESTRUCTURA DE LOS POLIOVIRUS; mayo, pág. 22.

Foley, James D. INTERFASES PARA ORDENADORES AVANZADOS; diciembre, pág. 76.

Fontes, Agustín G., Francisco F. de la Rosa, Miguel A. de la Rosa y Carlos Gómez-Moreno. PRODUCCIÓN FOTOQUÍMICA DE PERÓXIDO DE HIDRÓGENO; febrero, pág. 8.

Fox, Geoffrey C. y Paul C. Messina. ARQUITECTURAS AVANZADAS DE COMPUTADORES; diciembre, pág. 24.

Fraga, Serafín, Luis Seijo, Mercedes Campillo y Bhagirath Singh. PREVENCIÓN DEL RECHAZO DE TRASPLANTES; enero, pág. 6.

Francis, Peter y Stephen Self. HUNDIMIENTO DE LOS VOLCANES; agosto, pág. 54.

Frensley, W. R. TRANSISTORES DE ARSENIURO DE GALIO; octubre, pág. 50.

Gallo, Robert C. EL PRIMER RETROVIRUS HUMANO; febrero, pág. 44.

Gallo, Robert C. EL VIRUS DEL SIDA; marzo, pág. 30.

Gelernter, David. PROGRAMAS PARA COMPUTACIÓN AVANZADA; diciembre, pág. 52.

Gildea, Patricia, y George Miller. CÓMO APRENDEN LAS PALABRAS LOS NIÑOS; noviembre, pág. 80.

Giménez, Cecilio, y Federico Mayor. RECEPTORPATÍAS; marzo, pág. 10.

Glantz, Michael H. SEQUÍA EN ÁFRICA; agosto, pág. 26.

Gómez-Moreno, Carlos, Francisco F. de la Rosa, Miguel A. de la Rosa y Agustín G. Fontes. PRODUCCIÓN FOTOQUÍMICA DE PERÓXIDO DE HIDRÓGENO; febrero, pág. 8.

Gosselin, Edward A., y Lawrence S. Lerner. GALILEO Y EL FANTASMA DE BRUNO; enero, pág. 80.

Gould, James L., y P. Marler. Aprendizaje instintivo; marzo, pág. 46.
Habicht, Gail S., Gregory Beck y Jorge L. Benach. La enfermedad de lyme; septiembre, pág. 52.
Hamakawa, Yoshihiro. Energía fotovoltaica; junio, pág. 64.
Hasenstein, Karl-Keinz, Michael L. Evans y Randy Moore. Respuesta de las raíces a la gravedad; febrero, pág. 68.
Heiblum, Mordehai, y Lester F. Eastman. Electrones balísticos en semiconductores; abril, pág. 56.
Heinrich, Bernd. Termorregulación de las mariposas nocturnas de invierno; mayo, pág. 76.
Herrera, P., F. G. Bernáldez, C. Levassor, B. Peco y A. Sastre. Las aguas subterráneas en el paisaje; abril, pág. 8.
Hillis, W. Daniel. La máquina de conexión; agosto, pág. 72.
Hirsch, Martin S., y Joan C. Kaplan. Terapia antivírica; junio, pág. 50.
Hogle, James M., Marie Chow y David J. Filman. Estructura de los poliovirus; mayo, pág. 22.
Horton, Mark. Corredor swahili; noviembre, pág. 68.
Hut, Piet, y Gerald Jay Sussman. Computación avanzada en el campo de la ciencia; diciembre, pág. 96.
Ingersoll, Andrew P. Urano; marzo, pág. 20.
Jeffries, Andrew D., Peter R. Saulson, Robert E. Spero y Michael E. Zucker. Observatorios de ondas gravitatorias; agosto, pág. 44.
Jiménez, Fernando, y José A. Campos-Ortega. Bases de la neurogénesis en *Drosophila melanogaster*; septiembre, pág. 8.
Johnson, Kirk R., y C. Hans Nelson. Ballenas y morsas, labradoras del fondo marino; abril, pág. 68.
Johnson, Torrence V., Robert Hamilton Brown y Laurence A. Soderblom. Los satélites de Urano; junio, pág. 30.
Jones, Steven E., y Johann Rafelski. Fusión nuclear fría; septiembre, pág. 60.
Kahn, Robert E. Avances en redes de ordenadores; diciembre, pág. 86.
Kaplan, Joan C., y Martin S. Hirsch. Terapia antivírica; junio, pág. 50.
Koch, Christof, y Tomaso Poggio. Sinapsis que computan el movimiento; julio, pág. 28.
Krauss, Lawrence M. Materia oscura del universo; febrero, pág. 16.
Krisch, Alan D. Colisiones entre protones con espín; octubre, pág. 18.
Kryder, Mark H. Almacenamiento de datos para computación avanzada; diciembre, pág. 62.
Laeser, Richard P., William I. McLaughlin y Donna M. Wolff. Preparando el encuentro del *Voyager 2* con urano; enero, pág. 18.
Legge, Anthony J., y Peter A. Rowley-Conwy. Caza de gacelas en la siria de la edad de piedra; octubre, página 72.
Lerner, Lawrence S., y Edward A. Gosselin. Galileo y el fantasma de bruno; enero, pág. 80.
Letcher, Jr., John S., John K. Marshall, James C. Oliver III y Nils Salvesen. *Stars & stripes*; octubre, pág. 8.
Levassor, C., F. G. Bernáldez, P. Herrera, B. Peco y A. Sastre. Las aguas subterráneas en el paisaje; abril, pág. 8.
López Gorgé, Julio, M. Barón y A. Chueca. Inhibición de la fotosíntesis por herbicidas; julio, pág. 10.
Marler, Peter, y J. L. Gould. Aprendizaje instintivo; marzo, pág. 46.
Marrs, Barry L., y Douglas C. Youvan. Mecanismo molecular de la fotosíntesis; agosto, pág. 34.
Marshall, John K., John S. Letcher, Jr., James C. Oliver III y Nils Salvesen. *Stars & stripes*; octubre, pág. 8.
Masland, Richard H. Arquitectura funcional de la retina; febrero, pág. 56.
Mateu, Leonardo. La mielina; agosto, pág. 82.
Mathews, Grant F., y Victor E. Viola. Síntesis cósmica de litio, berilio y boro; julio, pág. 18.
Mathews, Max V., y John R. Pierce. El ordenador, instrumento musical; abril, pág. 84.
Mayor, Federico, y Cecilio Giménez. Receptorpatías; marzo, pág. 10.
McLaughlin, William I., Richard P. Laeser y Donna M. Wolff. Preparando el encuentro del *voyager 2* con urano; enero, pág. 18.
McMenamin, Mark A. S. La radiación de la vida animal; junio, pág. 72.
Metcalf, Harold J., y William D. Phillips. Enfriamiento y confinamiento de átomos; mayo, pág. 32.
Mezquita Pla, Jovita. El gen de la histona h1; junio, pág. 92.
Meindl, James D. Microcircuitos para computación avanzada; diciembre, pág. 34.
Messina, Paul C., y Geoffrey C. Fox. Arquitecturas avanzadas de computadores; diciembre, pág. 24.
Miller, George, y Patricia Gildea. Cómo aprenden las palabras niños; noviembre, pág. 80.
Mishkin, Mortimer, y Tim Appenzeller. Anatomía de la memoria; agosto, pág. 15.
Moore, Randy, Michael L. Evans y Karl-Keinz Hasenstein. Respuesta de las raíces a la gravedad; febrero, pág. 68.
Nelson, C. Hans, y Kirk R. Johnson. Ballenas y morsas, labradoras del fondo marino; abril, pág. 68.
Neves, Kenneth W., y Albert M. Erisman. Computación avanzada en la industria; diciembre, pág. 118.
Niklas, Karl J. Aerodinámica de la polinización eólica; septiembre, página 68.
Nossal, Sir Gustav, y Gordon L. Ada. La teoría de la selección clonal; octubre, pág. 40.
Oliver III, James C., John S. Letcher, Jr., John K. Marshall y N. Salvesen. *Stars & stripes*; octubre, pág. 8.
Oliver, José L., Manuel Ruiz Rejón y Carmelo Ruiz Rejón. Evolución de los cromosomas *b*; octubre, pág. 92.
Ostro, Marc J. Liposomas; marzo, página 74.
Packard, Norman H., James P. Crutchfield, J. Doyne Farmer y Robert S. Shaw. Caos; febrero, pág. 16.
Patel, C. Kumar N., y Nicolaas Bloembergen. Defensa estratégica y armas de energía dirigida; noviembre, pág. 14.
Patterson, David. Las causas del síndrome de down; octubre, pág. 28.
Peco, B., F. G. Bernáldez, P. Herrera, C. Levassor y A. Sastre. Las aguas subterráneas en el paisaje; abril, pág. 8.
Peled, A. La próxima revolución informática; diciembre, pág. 12.
Penney, Terry R., y Desikan Bharathan. Energía extraída del mar; marzo, pág. 58.
Petersen, Erik Brinch, y T. Douglas Price. Un campamento mesolítico en dinamarca; mayo, pág. 86.
Phillips, William D., y Harold J. Metcalf. Enfriamiento y confinamiento de átomos; mayo, pág. 32.
Pierce, John R., y Max V. Mathews. El ordenador, instrumento musical; abril, pág. 84.
Poggio, Tomaso, y Christof Koch. Sinapsis que computan el movimiento; julio, pág. 28.
Price, T. Douglas, y Erik Brinch Petersen. Un campamento mesolítico en dinamarca; mayo, pág. 86.
Psaltis, Demtri, y Yaser S. Abu-Mostafa. Computadoras opticoneuronales; mayo, pág. 58.
Rafelski, Johann, y Steven E. Jones. Fusión nuclear fría; septiembre, pág. 60.
Rennels, Glen D., y Edward H. Shortliffe. Computación avanzada en el campo de la medicina; diciembre, pág. 108.

Robertson, Douglas S., y William E. Carter. Geología por interferometría de muy larga base; enero, página 30.
Rosa, Francisco F. de la, Miguel A. de la Rosa, Agustín G. Fontes y Carlos Gómez-Moreno. Producción fotoquímica de peróxido de hidrógeno; febrero, pág. 8.
Rosa, Miguel A. de la, Francisco F. de la Rosa, Agustín G. Fontes y Carlos Gómez-Moreno. Producción fotoquímica de peróxido de hidrógeno; febrero, pág. 8.
Rowley-Conwy, Peter A., y Anthony J. Legge. Caza de gacelas en la Siria de la edad de piedra; octubre, página 72.
Ruiz Rejón, Carmelo, Manuel Ruiz Rejón y José L. Oliver. Evolución de los cromosomas *b*; octubre, página 92.
Ruiz Rejón, Manuel, Carmelo Ruiz Rejón y José L. Oliver. Evolución de los cromosomas *b*; octubre, página 92.
Ryder, Michael L. Evolución del vellón de la lana; marzo, pág. 84.
Salvesen, Nils, John S. Letcher, Jr., John K. Marshall y James C. Oliver III. *Stars & stripes*; octubre, pág. 8.
Sander, Leonard M. Crecimiento fractal; marzo, pág. 66.
Sastre, A., F. G. Bernáldez, P. Herrera, C. Levassor y B. Peco. Las aguas subterráneas en el paisaje; abril, pág. 8.
Saulson, Peter R., Andrew D. Jeffries, Robert E. Spero y Michael E. Zucker. Observatorios de ondas gravitatorias; agosto, pág. 44.
Schnapf, Julie L., y Denis A. Baylor. Respuesta de los fotorreceptores a la luz; junio, pág. 20.
Schneider, Stephen H. Modelos climáticos; julio, pág. 40.
Seijo, Luis, Serafín Fraga, Mercedes Campillo y Bhagirath Singh. Prevención del rechazo de trasplantes; enero, pág. 6.
Self, Stephen, y Peter Francis. Hundimiento de los volcanes; agosto, pág. 54.
Seu-Salerno, Martine, y J. Blakeway. El perfume; septiembre, pág. 38.
Shaham, Jacob. Los pulsares más antiguos del universo; abril, página 26.
Shaw, Robert S., James P. Crutchfield, J. Doyne Farmer y Norman H. Packard. Caos; febrero, pág. 16.
Shaw, Robert W. Contaminación atmosférica por partículas; octubre, pág. 82.
Shen, S. Acústica de las antiguas campanas chinas; junio, pág. 82.
Shortliffe, Edward H., y Glen D. Rennels. Computación avanzada en el campo de la medicina; diciembre, pág. 108.
Singh, Bhagirath, Serafín Fraga, Luis Seijo y Mercedes Campillo. Prevención del rechazo de trasplantes; enero, pág. 6.
Soderblom, Laurence A., Torrence V. Johnson y Robert Hamilton Brown. Los satélites de urano; junio, página 30.
Somero, George N., James J. Childress y H. Felbeck. Simbiosis en las profundidades marinas; julio, pág. 78.
Spero, Robert E., Andrew D. Jeffries, Peter R. Saulson y Michael E. Zucker. Observatorios de ondas gravitatorias; agosto, pág. 44.
Stahl, Franklin W. Recombinación genética; abril, pág. 42.
Stryer, Luber. Moléculas de excitación visual; septiembre, pág. 18.
Sussman, Gerald Jay, y Piet Hut. Computación avanzada en el campo de la ciencia; diciembre, pág. 96.
Taylor, Th. B. Armas nucleares de tercera generación; junio, pág. 8.
Treisman, Anne. Características y objetos del procesamiento visual; enero, pág. 68.
Varmus, Harold. Retrotranscripción; noviembre, pág. 32.
Vause, Chester A., y J. S. Walker. La reaparición de fases; julio, pág. 60.
Vellutino, Frank R. Dislexia; mayo, pág. 12.
Veltman, Martinus J. G. El bosón de higgs; enero, pág. 52.
Viola, Victor E., y Grant F. Mathews. Síntesis cósmica de litio, berilio y boro; julio, pág. 18.
Vlassara, Helen, Anthony Cerami y Michael Brownlee. Glucosa y envejecimiento; julio, pág. 52.
Walker, James S., y Ch. A. Vause. La reaparición de fases; julio, pág. 60.
Watts, Carol Martin, y Donald J. Watts. Un complejo de apartamentos romanos; febrero, pág. 84.
Watts, Donald J., y Carol Martin Watts. Un complejo de apartamentos romanos; febrero, pág. 84.
Wilson, David Gordon, Alec N. Brooks y Allan V. Abbott. Hidronaves de tracción humana; febrero, pág. 76.
Wolff, Donna M., Richard P. Laeser y William I. McLaughlin. Preparando el encuentro del *Voyager 2* con urano; enero, pág. 18.
Youvan, Douglas C., y Barry L. Marrs. Mecanismo molecular de la fotosíntesis; agosto, pág. 34.
Zapol, Warren M. Adaptaciones al buceo de la foca weddel; agosto, pág. 64.
Zucker, Michael E., Andrew D. Jeffries, Peter R. Saulson y Robert E. Spero. Observatorios de ondas gravitatorias; agosto, pág. 44.
Zuckerman, Harriet, y J. R. Cole. Rendimiento de la mujer casada en la investigación; abril, pág. 76.

TITULOS

Adaptaciones al buceo de la foca de weddel. Warren M. Zapol; agosto, pág. 64.
Acústica de las antiguas campanas chinas. Sinyan Shen; junio, pág. 82.
Aerodinámica de la polinización eólica. K. J. Niklas; septiembre, pág. 68.
Aguas subterráneas en el paisaje, las. F. G. Bernáldez, P. Herrera, C. Levassor, B. Peco y A. Sastre; abril, pág. 8.
Almacenamiento de datos para computación avanzada. Mark, H. Kryder; diciembre, pág. 62.
Anatomía de la memoria. Mortimer Mishkin y Tim Appenzeller; agosto, pág. 15.
Anillos de urano, los. Jeffrey N. Cuzzi y Larry W. Esposito; septiembre, pág. 28.
Aprendizaje instintivo. James L. Gould y P. Marler; marzo, pág. 46.
Arcos y bóvedas en el próximo oriente. Gus W. Van Beek; septiembre, pág. 76.
Armas nucleares de tercera generación. Theodore B. Taylor; junio, pág. 8.
arn, función enzimática del. Thomas R. Cech; enero, pág. 42.
Arquitectura funcional de la retina. Richard H. Masland; febrero, pág. 56.
Arquitecturas avanzadas de computadores. Geoffrey C. Fox y Paul C. Messina; diciembre, pág. 24.
Arseniuro de galio, transistores de. W. R. Frensley; octubre, pág. 50.
Atomos, enfriamiento y confinamiento de. William D. Phillips y Harold J. Metcalf; mayo, pág. 32.
Avances en redes de ordenadores. Robert E. Kahn; diciembre, pág. 86.
Ballenas y morsas, labradoras del fondo marino. C. Hans Nelson y Kirk R. Johnson; abril, pág. 68.
Bases de la neurogénesis en *Drosophila melanogaster*. José A. Campos-Ortega y Fernando Jiménez; septiembre, pág. 8.
Boson de Higgs, el. Martinus J.G. Veltman; enero, pág. 52.
Campamento mesolítico en Dinamarca, un. T. Douglas Price y Erik Brinch Petersen; mayo, pág. 86.

CAMPANAS CHINAS, ACÚSTICA DE LAS ANTIGUAS. Sinyan Shen; junio, página 82.
CAOS. James P. Crutchfield, J. Doyne, Norman H. Packard y Robert S. Shaw; febrero, pág. 16.
CARACTERÍSTICAS Y OBJETOS DEL PROCESAMIENTO VISUAL. Anne Treisman; enero, pág. 68.
CAUSAS DEL SÍNDROME DE DOWN, LAS. David Patterson; octubre, pág. 28.
CAZA DE GACELAS EN LA SIRIA DE LA EDAD DE PIEDRA. Anthony J. Legge y Peter A. Rowley-Conwy; octubre, pág. 72.
COLISIONES ENTRE PROTONES CON ESPÍN. Alan D. Krisch; octubre, pág. 18.
CÓMO APRENDEN LAS PALABRAS LOS NIÑOS. George Miller y Patricia Gildea; noviembre, pág. 80.
COMPLEJO DE APARTAMENTOS ROMANOS, UN. Donald J. Watts y Carol Martin Watts; febrero, pág. 84.
COMPUTACIÓN AVANZADA, ALMACENAMIENTO DE DATOS PARA. Mark H. Kryder; diciembre, pág. 62.
COMPUTACIÓN AVANZADA EN EL CAMPO DE LA CIENCIA. Piet Hut y Gerald Jay Sussman; diciembre, pág. 96.
COMPUTACIÓN AVANZADA EN LA INDUSTRIA. A. M. Erisman y K. W. Neves; diciembre, pág. 118.
COMPUTACIÓN AVANZADA EN EL CAMPO DE LA MEDICINA. Glen D. Rennels y Edward H. Shortliffe; diciembre, pág. 108.
COMPUTACIÓN AVANZADA, MICROCIRCUITOS PARA. James D. Meindl; diciembre, pág. 34.
COMPUTACIÓN AVANZADA, PROGRAMAS PARA. David Gelernter; diciembre, pág. 52.
COMPUTADORAS OPTICONEURONALES. Yaser S. Abu-Mostafa y Demtri Psaltis; mayo, pág. 58.
COMPUTADORES, ARQUITECTURAS AVANZADAS DE. Geoffrey C. Fox y Paul C. Messina; diciembre, pág. 24.
CONTAMINACIÓN ATMOSFÉRICA POR PARTÍCULAS. Robert W. Shaw; octubre, pág. 82.
CORREDOR SWAHILI. Mark Horton; noviembre, pág. 68.
CRECIMIENTO FRACTAL. Leonard M. Sander; marzo, pág. 66.
CROMOSOMAS *B*, EVOLUCIÓN DE LOS. Manuel Ruiz Rejón, Carmelo Ruiz Rejón y J. L. Oliver; octubre, pág. 92.
DEFENSA ESTRATÉGICA Y ARMAS DE ENERGÍA DIRIGIDA. C. Kumar N. Patel y Nicolaas Bloembergen; noviembre, pág. 14.
DETONACIÓN DE EXPLOSIVOS. William C. Davis; julio, pág. 70.
DISLEXIA. Frank R. Vellutino; mayo, pág. 12.
ELECTRONES BALÍSTICOS EN SEMICONDUCTORES. Mordehai Heiblum y Lester F. Eastman; abril, pàg. 56.
ELECTRUROS. James L. Dye; noviembre, pág. 44.
ENERGÍA EXTRAÍDA DEL MAR. T. R. Penney y D. Bharathan; marzo, pág. 58.
ENERGÍA FOTOVOLTAICA. Yoshiro Hamakawa; junio, pág. 64.
ENFERMEDAD DE LYME, LA. Gail S. Habicht, Gregory Beck y Jorge L. Benach; septiembre, pág. 52.
ENFRIAMIENTO Y CONFINAMIENTO DE ÁTOMOS. William D. Phillips y Harold J. Metcalf; mayo, pág. 32.
ESTRUCTURA DE LOS POLIOVIRUS. James H. Hogle, Marie Chow y David J. Filman; mayo, pág. 12.
EVOLUCIÓN DE LOS CROMOSOMAS *B*. Manuel Ruiz Rejón, Carmelo Ruiz Rejón y J. L. Oliver; octubre, pág. 92.
EVOLUCIÓN DEL VELLÓN DE LA LANA. Michael L. Ryder; marzo, pág. 84.
FIJACIÓN BIOLÓGICA DEL NITRÓGENO. Francisco Castillo y Jacobo Cárdenas; noviembre, pág. 88.
FOCA DE WEDDEL, ADAPTACIONES AL BUCEO DE LA. Warren M. Zapol; agosto, pág. 64.
FOTOSÍNTESIS, MECANISMO MOLECULAR DE LA. Douglas C. Youvan y Barry L. Marrs; agosto, pág. 34.
FUNCIÓN ENZIMÁTICA DEL ARN. Thomas R. Cech; enero, pág. 42.
FUSIÓN NUCLEAR FRÍA. J. Rafelski y Steven E. Jones; septiembre, pág. 60.
GALAXIAS, EL GRANDIOSO DESFILE DE LAS. A. Dressler; noviembre, pág. 22.
GALILEO Y EL FANTASMA DE BRUNO. Lawrence S. Lerner y Edward A. Gosselin; enero, pág. 80.
GEN DE LA HISTONA H1, EL. Jovita Mezquita Pla; junio, pág. 92.
GEOLOGÍA POR INTERFEROMETRÍA DE MUY LARGA BASE. W. E. Carter y Douglas S. Robertson; enero, pág. 30.
GLUCOSA Y ENVEJECIMIENTO. Anthony Cerami, Helen Vlassara y Michael Brownlee; julio, pág. 52.
GRANDIOSO DESFILE DE LAS GALAXIAS, EL. A. Dressler; noviembre, pág. 22.
HERBICIDAS, INHIBICIÓN DE LA FOTOSÍNTESIS POR. Matilde Barón, Ana Chueca y J. López Gorgé; julio, pág. 10.
HIDRONAVES DE TRACCIÓN HUMANA. Alec N. Brooks, Allan V. Abbott y D. Gordon Wilson; febrero, pág. 76.
HISTONA H1, EL GEN DE LA. Jovita Mezquita; junio, pág. 92.
HUNDIMIENTO DE LOS VOLCANES. Peter Francis y S. Self; agosto, pág. 54.
INFORMÁTICA, LA PRÓXIMA REVOLUCIÓN; diciembre, pág. 12.
INHIBICIÓN DE LA FOTOSÍNTESIS POR HERBICIDAS. Matilde Barón, Ana Chueca y Julio López Gorgé; julio, pág. 10.
INSTRUMENTO MUSICAL, EL ORDENADOR. Max V. Mathews y John R. Pierce; abril, pág. 84.
INTERFASES PARA ORDENADORES AVANZADOS. James D. Foley; diciembre, pág. 76.
INTERFEROMETRÍA DE MUY LARGA BASE, GEOLOGÍA POR. William E. Carter y D. S. Robertson; enero, pág. 30.
LANA, EVOLUCIÓN DEL VELLÓN DE LA. Michael L. Ryder; marzo, pág. 84.
LIPOSOMAS. M. J. Ostro; marzo, pág. 74.
LITIO, BERILIO Y BORO, SÍNTESIS CÓSMICA DE. Victor E. Viola y Grant F. Mathews; julio, pág. 18.
MÁQUINA DE CONEXIÓN, LA. W. Daniel Hillis; agosto, pág. 72.
MARIPOSAS NOCTURNAS DE INVIERNO, TERMORREGULACIÓN DE LAS. Bernd Heinrich; mayo, pág. 76.
MATERIA OSCURA DEL UNIVERSO. Lawrence M. Krauss; febrero, pág. 30.
MECANISMO MOLECULAR DE LA FOTOSÍNTESIS. Douglas C. Youvan y Barry L. Marrs; agosto, pág. 34.
MIMETISMO VEGETAL. Spencer C.H. Barrett; noviembre, pág. 58.
MEMORIA, ANATOMÍA DE LA. Mortimer Mishkin y Tim Appenzeller; agosto, pág. 15.
MICROCIRCUITOS PARA COMPUTACIÓN AVANZADA. James D. Meindl; diciembre, pág. 34.
MICROTÚBULO, MOTOR INTRACELULAR, EL. Robert Day Allen; abril, pág. 18.
MIELINA, LA. Leonardo Mateu; agosto, pág. 82.
MODELOS CLIMÁTICOS. Stephen H. Schneider; julio, pág. 40.
MOLÉCULAS DE EXCITACIÓN VISUAL. Luber Stryer; septiembre, pág. 18.
MORSAS Y BALLENAS, LABRADORAS DEL FONDO MARINO. C. Hans Nelson y Kirk R. Johnson; abril, pág. 68.
NEUROGÉNESIS EN *DROSOPHILA MELANOGASTER*, BASES DE LA. José A. Campos-Ortega y Fernando Jiménez; septiembre, pág. 8.
NIÑO Y LA OSCILACIÓN DEL SUR, EL. Francisco Chávez; mayo, pág. 46.
OBSERVATORIOS DE ONDAS GRAVITATORIAS. Andrew D. Jeffries, Peter R. Saulson, Robert E. Spero y Michael E. Zucker; agosto, pág. 44.
ORDENADOR, INSTRUMENTO MUSICAL, EL. Max V. Mathews y John R. Pierce; abril, pág. 84.
ORDENADORES AVANZADOS, INTERFASES PARA. James D. Foley; diciembre, pág. 76.
OSCILACIÓN DEL SUR, EL NIÑO Y LA. Francisco Chávez; mayo, pág. 46.
PECES ANTÁRTICOS. Joseph T. Eastman y Arthur L. DeVries; enero, pág. 60.
PERFUME, EL. Martine Seu-Salerno y John Blakeway; septiembre, pág. 38.

PERÓXIDO DE HIDRÓGENO, PRODUCCIÓN FOTOQUÍMICA DE. Francisco F. de la Rosa, Miguel A. de la Rosa, Agustín C. Fontes y Carlos Gómez-Moreno, febrero, pág. 8.
POLINIZACIÓN EÓLICA, AERODINÁMICA DE LA. Karl J. Niklas; septiembre, página 68.
POLIOVIRUS, ESTRUCTURA DE LOS. James H. Hogle, Marie Chow y David J. Filman; mayo, pág. 22.
PREPARANDO EL ENCUENTRO DEL *VOYAGER 2* CON URANO. Richard P. Laeser, William I. McLaughlin y Donna M. Wolff; enero, pág. 18.
PREVENCIÓN DEL RECHAZO DE TRASPLANTES. Serafín Fraga, Luis Seijo, Mercedes Campillo y Bhagirath Singh; enero, pág. 6.
PRIMER RETROVIRUS HUMANO, EL. Robert C. Gallo; febrero, pág. 44.
PROCESAMIENTO VISUAL, CARACTERÍSTICAS Y OBJETOS DEL PROCESAMIENTO VISUAL. Anne Treisman; enero, pág. 68.
PRODUCCIÓN FOTOQUÍMICA DEL PERÓXIDO DE HIDRÓGENO. Francisco F. de la Rosa, Miguel A. de la Rosa, Agustín G. Fontes y Carlos Gómez-Moreno; febrero, pág. 8.
PROGRAMAS PARA COMPUTACIÓN AVANZADA. David Gelernter; diciembre, pág. 52.
PRÓXIMA REVOLUCIÓN INFORMÁTICA, LA. Abraham Peled; diciembre, pág. 12.
PULSARES MÁS ANTIGUOS DEL UNIVERSO, LOS. Jacob Shaham; abril, pág. 26.
RADIACIÓN DE LA VIDA ANIMAL, LA. Mark A. S. McMenamin; junio, página 72.
REAPARICIÓN DE FASES, LA. James S. Walker y Chester A. Vause; julio, pág. 60.
RECEPTORPATÍAS. Federico Mayor y Cecilio Giménez; marzo, pág. 10.
RECOMBINACIÓN GENÉTICA. Franklin W. Sthal; abril, pág. 42.
REDES DE ORDENADORES, AVANCES EN. Robert E. Kahn; diciembre, pág. 86.
RENDIMIENTO DE LA MUJER CASADA EN LA INVESTIGACIÓN. Jonathan R. Cole y Harriet Zuckerman; abril, pág. 76.
RESPUESTA DE LAS RAÍCES A LA GRAVEDAD. Michael L. Evans, Randy Moore y Karl-Keinz Hasenstein; febrero, pág. 68.
RESPUESTA DE LOS FOTORRECEPTORES A LA LUZ. Julie L. Schnapf y Denis A. Baylor; junio, pág. 20.
RETINA, ARQUITECTURA FUNCIONAL DE LA. Richard H. Masland; febrero, pág. 56.
RETROTRANSCRIPCIÓN. Harold Varmus; noviembre, pág. 32.
RETROVIRUS HUMANO, EL PRIMER. Robert C. Gallo; febrero, pág. 44.
ROTURA DE LOS CONTINENTES, LA. Enrico Bonatti; mayo, pág. 66.
SATÉLITES DE URANO, LOS. Torrence V. Johnson, Robert Hamilton Brown y L. A. Soderblom; junio, pág. 30.
SELECCIÓN CLONAL, LA TEORÍA DE LA. Gordon L. Ada y Sir Gustav Nossal; octubre, pág. 40.
SEMICONDUCTORES, ELECTRONES BALÍSTICOS EN. Mordehai Heiblum y Lester F. Eastman; abril, pág. 56.
SEQUÍA EN ÁFRICA. Michael H. Glantz; agosto, pág. 26.
SIMBIOSIS EN LAS PROFUNDIDADES MARINAS. James J. Childress, Horts Felbeck y G. N. Somero; julio, pág. 78.
SINAPSIS QUE COMPUTAN EL MOVIMIENTO. Tomaso Poggio y Christof Koch; julio, pág. 28.
SÍNDROME DE DOWN, LAS CAUSAS DEL. David Patterson; octubre, pág. 28.
SÍNTESIS CÓSMICA DE LITIO, BERILIO Y BORO. Víctor E. Viola y Grant F. Mathews; julio, pág. 18.
STARS & STRIPES. John S. Letcher, Jr., John K. Marshall, James C. Oliver III y Nils Salvesen; octubre, pág. 8.
TEORÍA DE LA SELECCIÓN CLONAL, LA. Gordon L. Ada y Sir Gustav Nossal; octubre, pág. 40.
TERAPIA ANTIVÍRICA. Martins Hirsch y Joan C. Kaplan; junio, pág. 50.
TERMORREGULACIÓN DE LAS MARIPOSAS NOCTURNAS DE INVIERNO. Bernd Heinrich; mayo, pág. 76.
TRANSISTORES DE ARSENIURO DE GALIO. W. R. Frensley; octubre, pág. 50.
TRASPLANTES, PREVENCIÓN DEL RECHAZO DE. Serafín Fraga, Luis Seijo, Mercedes Campillo y Bhagirath Singh; enero, pág. 6.
URANO. Andrew P. Ingersoll; marzo, pág. 20.
URANO, LOS ANILLOS DE. Jeffrey N. Cuzzi y Larry W. Esposito; septiembre, pág. 28.
URANO, PREPARANDO EL ENCUENTRO DEL *VOYAGER 2* CON. Richard P. Laeser, William I. McLaughlin y Donna M. Wolff; enero, pág. 18.
URANO, SATÉLITES DE. Torrence V. Johnson, R. Hamilton Brown y Laurence A. Soderblom; junio, pág. 30.
VIRUS DEL SIDA, EL. Robert C. Gallo; marzo, pág. 30.
VOLCANES, HUNDIMIENTO DE LOS. Peter Francis y S. Self; agosto, pág. 54.

JUEGOS DE ORDENADOR

Autor: A.K. Dewdney

A los oídos de algunos, los sonidos de la informática parecen música; junio, pág. 102.
"After MAD": Juego computarizado de estrategia nuclear que acaba en un "dilema del prisionero"; diciembre, pág. 128.
Algoblemas: problemas donde el pensamiento avanza hacia la solución por vías algorítmicas; agosto, página 94.
Escalas verbales y una Torre de Babel llevan a alturas computacionales casi imposibles de asaltar, Las; octubre, pág. 104.
Indagando en la extraña atracción del caos físico; septiembre, pág. 86.
Juego Vida cuenta ya con sucesores en tres dimensiones, El; abril, pág. 94.
Montañas fractales, plantas graftales y múltiples ordinográficos en Pixar; febrero, pág. 92.
Personalidades diversas tratan de hallar equilibrio social en una fiesta informática; noviembre, pág. 98.
Programa de RATONES va royendo paso a paso hasta la victoria en el primer torneo de Guerra Nuclear, El; marzo, pág. 94.
Sobre toros, valores y programas de parquet; julio, pág. 86.
Star Treck emerge de las cavernas y se instala en los videojuegos domésticos; enero, pág. 89.
Sutilezas de los anillos mágicos, último hallazgo de Ernö Rubik, genial inventor del cubo diabólico; mayo, pág. 96.

TALLER Y LABORATORIO

Autor: Jearl Walker

Cómo calcular la distancia de un punto al Sol observando la trayectoria de un meteorito; mayo, pág. 102.
Cómo construir un barómetro con agua en vez de mercurio; junio, pág. 107.
Cómo cruzar un laberinto sin perderse ni aturdirse; febrero, pág. 98.
Hablemos de desapariciones, incluida la del evanescente Gato de Cheshire; julio, pág. 91.
Hay sustancias fibrosas y pegajosas que propenden a formar rosarios de gotas; noviembre, pág. 104.
Hiperoscopio y el pseudoscopio facilitan la experimentación sobre la visión tridimensional, El; enero, página 94.
Música y vapores de amoníaco excitan los colores de las películas de jabón; octubre, pág. 110.
Por qué un fluido acelera su curso por el interior de un tubo pinzado; septiembre, pág. 92.
Reflejos del agua y sus curiosas propiedades, Los; marzo, pág. 100.
Rompecabezas en dos y tres dimensiones y las reglas para simplificar su solución; agosto, pág. 98.
Se revela el secreto de la rapidez de cocción de los hornos microondas; abril, pág. 100.